```
AAE06229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      January 20, 2004, 18:05:34; Search time 44.971 Seconds (without alignments) 102.356 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
2. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
3. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
4. (SIDSI) gogdata/geneseqy-embl/AA1981.DAT:*
5. (SIDSI) gogdata/geneseqp-embl/AA1981.DAT:*
6. (SIDSI) gogdata/geneseqp-embl/AA1981.DAT:*
7. (SIDSI) gogdata/geneseqp-embl/AA1986.DAT:*
8. (SIDSI) gogdata/geneseqp-embl/AA1981.DAT:*
8. (SIDSI) gogdata/geneseqp-embl/AA1980.DAT:*
9. (SIDSI) gogdata/geneseqp-embl/AA1980.DAT:*
10. (SIDSI) gogdata/geneseqp-embl/AA1980.DAT:*
11. (SIDSI) gogdata/geneseqp-embl/AA1991.DAT:*
12. (SIDSI) gogdata/geneseqp-embl/AA1992.DAT:*
13. (SIDSI) gogdata/geneseqp-embl/AA1992.DAT:*
14. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA1992.DAT:*
15. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA1992.DAT:*
16. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA1995.DAT:*
17. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA1995.DAT:*
18. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA1995.DAT:*
18. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA1999.DAT:*
18. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA1999.DAT:*
20. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA1999.DAT:*
21. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA1999.DAT:*
22. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA1999.DAT:*
23. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA2001.DAT:*
24. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA2001.DAT:*
23. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA2001.DAT:*
24. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA2001.DAT:*
24. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA2001.DAT:*
25. (SIDSI) gogdata/geneseqfgeneseqp-embl/AA2001.DAT:*
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                               US-10-018-103A-4
183
1 KHKHKHKHKGKHKHKHKKKK 29
                                                                                                                                                                                                                                                                                                                                                                                                           Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                      1107863 seqs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                  OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                    BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A_Geneseq_19Jun03:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Title:
Perfect score:
                                                                                                                                                                                                                                                                                                    Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ••
                                                                                                                                                                                                                                                             Sequence:
                                                                                                                                                                                                                                                                                                                                                                      Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Database
                                                                                                                             Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description		Histidine copolyme	H-K2b polymer. Un	H-K3b polymer. Un	H-K4b polymer. Un	Histidine copolyme	Histidine copolyme	Histidine copolyme	Histidine copolyme	H influenzae BVH-N
SUMMAKIES	ΙD		AAE06229	AAE06242	AAE06243	AAE06244	AAE06228	AAE06234	AAE06241	AAE06232	AA017803
	DB	!	22	22	22	22	22	22	22	22	23
	* Query e Match Length DB ID		58	39	59	79	19	21	. 21	53	335
	Query Match		100.0	92.1	92.1	86.3	65.6	9.59	65.6	63.9	62.8
	Score		183	168.5	168.5	158	120	120	120	117	115
	Result No.		-	8	m	4	S	9	7	œ	σ

Haemophilus influe	Human ovarian canc	Human protein sequ	Human protein segu	Human CIF150/hTAF1	Human CIF150/hTAFI	Propionibacterium	HH-K3b polymer. U	HH-K4b polymer. U	S. cerevisiae BAX-	K2b poly	Sequence of Histid	Histidine copolyme	Drosophila melanog	Drosophila G-prote	Arabidopsis thalia	Arabidopsis thalia	Arabidopsis thalia	Histidine copolyme	Human polypeptide	Drosophila melanog	Human colon cancer	Human protein sequ	Human polypeptide	Hypoxia-repressed	Human secreted pro	Human liver peptid	Human peptide #14	#15	Protein #14 encode	Human brain expres	Human bone marrow	a	Peptide #15 encode	e #15	Human peptide enco
AAW61155	ABG96401	AAB93454	AAB93135	AAY29083	AAY44263	AAU44422	AAE06246	AAE06247	ABG93037	AAE06245	AAR24393	AAE06227	ABB63629	AAU38954	AAG31823	AAG31822	AAG31821	AAE06238	AA010602	ABB68805	AAG75565	AAB92603	AAM40326	ABP65080	AAY07923	ABG47362	ABB27363	ABB32509	ABB18015	AAM53343	AAM65721	AAM13580	AAM25978	AAM01333	ABG35350
13	23	22	22	20	21	22	22	22	23	22	13	22	22	22	21	21	21	22	22	22	22	22	22	23	20	22	22	22	22	22	22	22	22	22	23
337	1043	271	504	1199	1199	129	62	83	980	41	351	15	337	337	359	375	385	19	81	639	106	483	931	931	78	82	82	82	82	82	82	85	82	82	82
62.3	60.4	57.1	57.1	57.1	57.1	53.6	52.7	52.7	52.5	52.2	51.9	51.4	50.8	50.8	49.7	49.7	49.7	49.2	48.6	48.4	47.0	47.0	47.0	47.0	45.9	45.9	45.9	45.9	45.9	45.9	45.9	45.9	45.9	45.9	45.9
114	110.5	104.5	104.5	104.5	104.5	96	9	96.5	96	95.5	92	94	93	93	91	91	91	06	89	88.5	98	98	98	98	84	84	84	84	84	84	84	84	84	84	84
10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

Histidine copolymer, H-K 29-mer peptide AAE06229 standard; peptide; 29 AA. 25-SEP-2001 (first entry) AAE06229;

Histidine copolymer; enzyme deficiency; phenylalanine hydroxylase; ornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; purine nucleoside phosphorylase; alphal-antitrypsin; apolipoprotein E; chronic granulomatous disease; sickle cell; beta-thalassemia; anaemia; familial hypercholegerolaemia; low-density lipoprotein receptor; phenylketonuria; Faconi's anaemia; low-density lipoprotein receptor; phenylketonuria; Faconi's anaemia; haemophilia; muscular dystrophy; cystic fibrosis; Parkinson's disease; retinitis pigmentosa; cytostatic; lysosomal storage disease; mycoolysaccharide type I disease; cardiant; diabetic retinopathy; human immunodeficiency virus disease; infection; cardiac disease; peripheral vascular disease; antisickling; arthritis; nootropic; haemostatic; virucide; gene therapy.

Unidentified

WO200147496-A1 05-JUL-2001 20-DEC-2000; 2000WO-US34603.

99US-0173576. 29-DEC-1999; 20-DEC-2000; 2000WO-US34603

WO200147496-A1

05-JUL-2001.

99US-0173576

29-DEC-1999;

(MIXS/) MIXSON A J.

Mixson AJ;

WPI; 2001-425579/45.

```
The invention relates to a pharmaceutical agent delivery composition comprising a transport polymer comprising a linear or branched peptide having about 1 to to 300 anino acid residues, having about 5 to 100% thatidine residues, and optionally having 1 to 95% non-histidine residues, at least one pharmaceutical agent and optionally one or more intracellular delivery components in association with the transport collymer. The pharmaceutical agent delivery composition is useful for a tollywering a pharmaceutical agent delivery composition is useful for delivering a pharmaceutical agent to the interior of a cell and is useful for used in gene therapy for treating adenosine deaminase deficiency, purine concleoside phosphorylase deficiency, chronic granulomatous disease with defective p47phox, sickle cell with HBS, beta-thalassaemia due to inadequate production of beta-haemoglobin, familial hypercholesterolaemia con a defective low-density lipoprotein receptor, phonylactonine deficiency, alphal-antitrypsin deficiency, ornithine transcarbamylase deficiency, Paconi's anaemia, apolipoprotein E deficiency, haemophila A and B due to factor VIII and IX deficiency, cornithine transcarbamylase deficiency, retinitis pigmentosa, lysosomal storage tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage cursuina, cystic fibrosis due to CFTR mutations, Parkinson due to tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage cinfection, acquired anaemia, cardiac and peripheral vascular disease infection, acquired anaemia, cardiac and peripheral vascular disease, virus and arthritis. The present peptide sequence is histidine copolymer.
                                                                                                                                              Pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues
                                                                                                                                                                                                                                                                                  Claim 3; Page 36; 64pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                   WPI; 2001-425579/45.
(MIXS/) MIXSON A J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        29 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    H-K2b polymer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                     Mixson AJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAE06242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 2
  d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
```

```
Gaps
                                         ö
 Length 29;
                                         Indels
'Match 100.0%; Score 183; DB 22; Local Similarity 100.0%; Pred. No. 5.7e-17; es 29; Conservative 0; Mismatches 0;
                                                                                                     1 KHKHKHKHKGKHKHKHKHKHKHK 29
                                                                               1 КНКНКНКНКСКНКНКНКНКНКСКНКНК 29
```

ornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; purine nucleoside phosphorylase; alphal-antitrypsin; apolipoprotein E; chronic granulomatous disease; alckle cell; beta-thalassaemia annemia; familial hypercholesterolaemia; low-density lipoprotein receptor; phenylketonuria; Faconi's annemia; haemophilia; muscular dystrophy; cystic fibrosis; Parkinson's disease; retinitis pigmentosa; cytostatic; lysosomal storage disease; mycopolysaccharide type I disease; cardiant; diabetic retinopathy; human immunodeficiency virus disease; infection; cardiac disease; partipheral vascular disease; antisickling; arthritis; nootropic; haemostatic; virucide; gene therapy. Enzyme deficiency; phenylalanine hydroxylase; AAE06242 standard; peptide; 39 AA

```
Unidentified
```

```
the invention transport polymer comprising a linear or branched peptide

Comprising a transport polymer comprising a linear or branched peptide

Comprising a transport polymer comprising a linear or branched peptide

Confistidine residues, and optionally having it of 9% non-histidine

cresidues, at least one pharmaceutical agent and optionally one or more

intracellular delivery components in association with the transport

Conjumer. The pharmaceutical agent delivery composition is useful for

closed in gene therapy for treating adenosine deaminase deficiency, purine

cused in gene therapy for treating adenosine deaminase deficiency, purine

cused in gene therapy for treating adenosine deaminase deficiency, purine

cused in gene therapy for treating adenosine deaminase deficiency, purine

cused in gene therapy for treating adenosine deaminase deficiency, purine

cused in gene to peta-haemoglobin, familial hypercholesterolaemia

continue production of beta-haemoglobin, familial hypercholesterolaemia

continue adefective low-density lipoprotein receptor, phenylketouria due

contituine transcarbamylase deficiency, alphal-antitrypsin deficiency,

confitchine transcarbamylase deficiency, alphal-antitrypsin deficiency,

confitchine transcarbamylase deficiency, recentis anaemia, apolipoprotein

confitchine transcarbamylase deficiency, retinitis pignentosa, lysosomal storage

context on a servedly perceptor of the partine pignentosa, lysosomal storage

confitchine transcarbamylase deficiency, retinitis pignentosa, lysosomal storage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              disease (i.e., mycopolysacchari, type 1, Hunter, Hurler and Gaucher), disbetic retinopathy, human immunodeficiency virus disease, virus infection, acquired anaemia, cardiac and peripheral vascular disease and arthritis. The present peptide sequence is H-K2b polymer.
                                                                                                                                                                                                                                                                                                                                             Pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              invention relates to a pharmaceutical agent delivery composition
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 168.5; DB 22; Length 39;
Pred. No. 5.3e-15;
0; Mismatches 0; Indels 9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Example; Page 23; 64pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 76.3
nes 29; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      39 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best
à
```

1 КНКНКНККК-----КНКНКНКНККККККК 29 AAE06243 standard; peptide; 59 (first entry) H-K3b polymer. 25-SEP-2001 RESULT 3 AAE06243 셤

ornithine transcarbamylaes; adenosine deaminase; tyrosine hydroxylase; purine nucleoside phosphorylase; alphal-antitrypsin; apolipoprotein E; chronic granulomatous disease; sickle cell; beta-thalassemila; anaemia; familial hypercholesterolaemia; low-density lipoprotein receptor; phenylketonuria; Faconi's anaemia; haemophilia; muscular dystrophy; cystic fibrosis; Parkinson's disease; retinitis pigmentosa; cytostatic;

Enzyme deficiency; phenylalanine hydroxylase;

ä

ornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; purine nucleoside phosphorylase; alphal-antitrypsin; apolipoprotein B; chronic granulomatous disease; sickle cell; beta-thalassaemia; anaemia; familial hypercholesterolaemia; low-density lipoprotein receptor; phenylketomuria; Raconi's anaemia; haemophilia; muscular dystrophy; cystic fibrosis; Parkinson's disease; retinitis pigmentosa; cytostatic;

Enzyme deficiency; phenylalanine hydroxylase;

lysosomal storage disease; mycopolysaccharide type I disease; cardiant; diabetic retinopathy; human immunodeficiency virus disease; infection; cardiac disease; peripheral vascular disease; antisickling; arthritis;

nootropic; haemostatic; virucide; gene therapy.

20-DEC-2000; 2000WO-US34603.

WO200147496-A1. Unidentified

05-JUL-2001

99US-0173576

29-DEC-1999;

(MIXS/) MIXSON A J.

```
comprising a transport bolymer comprising a linear or branched populate having about 10 to 300 amino acid residues, having about 5 to 100% institution and optionally having about 5 to 100% thistidine residues, at least one pharmaceutical agent and optionally one or more intracellular delivery components in association with the transport polymer. The pharmaceutical agent delivery composition is useful for delivering a pharmaceutical agent to the interior of a cell and is used in gene therapy for treating adenosine deaminase deficiency, purine uncleoside phosphorylase deficiency, chronic granulomatous disease with defective payphox, sickle cell with HbS, beta-thalassaemia due to inadequate production of beta-haemoglobin, familial hypercholesterolaemia cut a defective low-density lipoprotein receptor, phenyletronuira due to phenylalanine hydroxylase deficiency, raconi's anaemia, apolipoprotein confithine transcarbamylase deficiency, raconi's anaemia, apolipoprotein mucaular dystrophy due to dystrophin, laminin-2, or sacroglycans mutations, cystic fibrosis due to CFTR mutations, Parkinson due to typosine hydroxylase deficiency, retinitis pignmentosa, lysosomal storage deficiency, retinitis pignmentosa, lysosomal storage typosine hydroxylase deficiency, retinitis pignmentosa, lysosomal storage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
cardiant;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              disease (i.e., mycopolysaccharide type 1, Hunter, Hurler and Gaucher), diabetic retinopathy, human immunodeficiency virus disease, virus infection, acquired anaemia, cardiac and peripheral vascular disease and arthritis. The present peptide sequence is H-K3b polymer.
                diabetic retinopathy; human immunodeficiency virus disease; infection; cardiac disease; peripheral vascular disease; antisickling; arthritis; nootropic; haemostatic; virucide; gene therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pharmaceutical composition useful for delivering therapeutic agent for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               invention relates to a pharmaceutical agent delivery composition
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
lysosomal storage disease; mycopolysaccharide type 1 disease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 168.5; DB 22; Length 59;
Pred. No. 8.1e-15;
0; Mismatches 0; Indels 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 KHKHKHKHKG------KHKHKHKHKKKKHKHK 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Example; Page 23; 64pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          92.1%;
                                                                                                                                                                                                                                             20-DEC-2000; 2000WO-US34603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            76.3%;
                                                                                                                                                                                                                                                                                       99US-0173576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2001-425579/45.
                                                                                                                                                                                                                                                                                                                                   (MIXS/) MIXSON A J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity
hes 29; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             59 AA;
                                                                                                                                                  WO200147496-A1
                                                                                                                                                                                                                                                                                       29-DEC-1999;
                                                                                                           Unidentified
                                                                                                                                                                                              05-JUL-2001
                                                                                                                                                                                                                                                                                                                                                                                  Mixson AJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Loca
Matches
ઠ
```

Pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues

WPI; 2001-425579/45.

Mixson AJ;

Example; Page 23; 64pp; English.

```
The inversion relates to a pharmaceutical agent using the function comparising a transport bolymar comprising a linear or branched peptide having about 10 to 300 amino acid residues, having about 5 to 100% institution residues, and optionally having 1 to 95% anon-histidine residues, and optionally having about 5 to 100% intracellular delivery components in association with the transport or polymer. The pharmaceutical agent delivery composition is useful for delivering a pharmaceutical agent delivery composition is useful for delivering a pharmaceutical agent to the interior of a cell and is used in gene therapy for treating adenosine deaminase deficiency, purine uncleoside phosphorylase deficiency, chronic granulomatous disease with defective py7phox, sickle cell with HbS, beta-thalassaemia due to inadequate production of beta-haemoglobin, familial hypercholesterolaemia continue transcarbamylase deficiency, alphal-antitrypsin deficiency, crnithine transcarbamylase deficiency, Paconi's anaemia, apolipoprotein wiscular dystrophy due to dystrophin, laminin-2, or sacroglycans mutations, cystic fibrosis due to CPTR mutations, Parkinson due to tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage cire, mycopolypaccharide type 1, Hunter, Hurler and Gaucher), diabetic retinopathy, human immunodeficiency virus disease, virus infection, acquired anaemia, cardiac and peripheral vascular disease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The invention relates to a pharmaceutical agent delivery composition
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                infection, acquired anaemia, cardiac and peripheral vascular and arthritis. The present peptide sequence is H-K4b polymer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 79;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 KHKHKHKH-KGKHKHKHKHKHKH 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 158; DB 22;
Pred. No. 2.3e-13;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAE06228 standard; peptide; 19 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         86.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          74.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity 74.4
les 29; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     79 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

AAE06244 standard; peptide; 79 AA

RESULT 4 **AAE0624**4

셤

(first entry)

25-SEP-2001

AAE06244;

H-K4b polymer

5

AAE06228;

25-SEP-2001 (first entry)

Histidine copolymer, H-K 19-mer peptide.

Histidine copolymer; enzyme deficiency; phenylalanine hydroxylase; ornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; purine nucleoside phosphorylase; alphal-antitrypsin; apolipoprotein E; chronic granulomatous disease; slickle cell; beta-thalassaemia; anaemia; familial hypercholesterolaemia; low-density lipoprotein receptor; phenylketonuria; Paconi's anaemia; haemophilia; muscular dystrophy; cystic fibrosis; Parkinson's disease; retinitis pigmentosa; cytostatic; lysosomal storage disease; mycopolysaccharide type I disease; cardiant; diabetic retinopathy; human immunodeficiency virus disease; infection; cardiac disease; peripheral vascular disease; antisickling; arthritis; nootropic; haemostatic; virucide; gene therapy.

Unidentified.

WO200147496-A1

05-JUL-2001

20-DEC-2000; 2000WO-US34603.

99US-0173576. 29-DEC-1999;

(MIXS/) MIXSON A J.

Mixson AJ;

WPI; 2001-425579/45.

Pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues

Claim 3; Page 36; 64pp; English.

The invention relates to a pharmaceutical agent delivery composition comprising a transport polymer comprising a linear or branched peptide comprising about 10 to 300 amino acid residues, having about 5 to 100% considered and parting about 5 to 100% and optionally having 1 to 95% non-histidine residues, and optionally having 1 to 95% non-histidine credidues, at least one pharmaceutical agent and optionally one or more intracellular delivery components in association with the transport collymer. The pharmaceutical agent delivery composition is useful for delivering a pharmaceutical agent delivery composition is useful for used in gene therapy for treating adenosine deaminase deficiency, purine undecaside phosphorylase deficiency, chronic granulomatous disease with defective p47phox, sickle cell with HBS, beta-thalassaema due to inadequate production of beta-haemoglobin, familial hypercholesterolaemia con the defective p47phox, sickle cell with HBS, beta-thalassaema disease with the transcarbamylase deficiency, Paconi's anaemia, apolipoprotein contribine transcarbamylase deficiency, Paconi's anaemia, apolipoprotein muscular dystrophy due to dystrophin, lamini-2, or sacroglycans customations, cystic fibrosis due to CFTR mutations, Parkinson due to tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage contentions, cystic fibrosis due to CFTR mutations, Parkinson due to tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage confibency in munodofficiency virus disease, virus infection, acquired anaemia, cardiac and peripheral vascular disease confibency membritis. The present peptide sequence is histidine copolymer. H-K 19-mer

19 AA; Sequence

ö Gaps ö Length 19; Indels 65.6%; Score 120; DB 22; 100.0%; Pred. No. 3.7e-09; cive 0; Mismatches 0; Query Match Best Local Similarity 100. Matches 19; Conservative

книнининиемник 19 ٦ à 셤

RESULT 6

AAE06234 standard; peptide; 21 AA.

AAE06234;

25-SEP-2001 (first entry)

Histidine copolymer, K-HK peptide.

Histidine copolymer; enzyme deficiency; phenylalanine hydroxylase; ornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; purine nucleoside phosphorylase; alphal-antitrypsin; apolipoprotein B; chronic granulomatous disease; sichle cell; beta-thalassaemia; anaemia; familial hypercholesterolaemia; low-density lipoprotein receptor; phenylketonuria; Faconi's anaemia; haemophilia; muscular dystrophy; cystic fibrosis; Parkinson's disease; retinitis pigmentosa; cytostatic; lysosomal storage disease; mycopolysaccharide type I disease; cardiant; dabatic retinopathy; human immunodeficiency virus disease; cardiant; cardiant; dabatic disease; peripheral vascular disease; infection; cardiac disease; peripheral vascular disease; antisickling; arthritis; nootropic; haemostatic; virucide; gene therapy.

Unidentified.

WO200147496-A1.

05-JUL-2001.

20-DEC-2000; 2000WO-US34603.

29-DEC-1999; 99US-0173576.

(MIXS/) MIXSON A J.

Mixson AJ;

WPI; 2001-425579/45.

Pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues

Example, Page 23; 64pp; English.

The invention relates to a pharmaceutical agent delivery composition comprising a transport polymer comprising a linear or branched peptide comprising a transport polymer comprising a linear or branched peptide having about 10 to 300 amino acid residues, having about 5 to 100% comprising a taleast one pharmaceutical agent and optionally one or more intracellular delivery components in association with the transport complying a pharmaceutical agent delivery composition is useful for delivering a pharmaceutical agent delivery composition is useful for cased in gene therapy for treating adenosine deaminase deficiency, purine used in gene therapy for treating adenosine deaminase deficiency, purine cused in gene therapy for treating adenosine deaminase deficiency, put of defective p47phox, sickle cell with HbS, beta-thalassaemia due to inadequate production of beta-haemoglobin, familial hypercholesterolaemia compensional pharmaceutical selectioncy, alphal-antitrypsin deficiency, contithine transcarbamylase deficiency, Paconi's anaemia, apolipoprotein E deficiency, haemophilia A and B due to factor VIII and IX deficiency, mutations, cystic fibrosis due to GFTR mutations, Parkinson due to tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage contineations, cystic fibrosis due to GFTR mutations, Parkinson due to tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage contibettion, acquired anaemia, cardiac and peripheral vascular disease infection, acquired anaemia, cardiac and peripheral vascular disease and arthritis. The present peptide sequence is histidine copolymer

ខ្លួន្តប្

ઠે 셤

```
RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                          AAE06232
  888888888888
                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The invention relates to a pharmaceutical agent delivery composition comprising a transport polymer comprising a linear or branched peptide having about 10 to 300 amino acid residues, having about 5 to 100% histidine residues, and optionally having 1 to 95% non-histidine residues, at least one pharmaceutical agent and optionally one or more intracellular delivery components in association with the transport polymer. The pharmaceutical agent to the interior of a cell and is useful for delivering a pharmaceutical agent to the interior of a cell and is useful for delivering a pharmaceutical agent to the interior of a cell and is used thorapy for treating adenosine deaminase deficiency, purine nucleoside phosphorylase deficiency, chronic granulomatous disease with defective pfyphox, sickle cell with HbS, beta-chalassaemia due to inadequate production of beta-haemoglobin, familial hypercholesterolaemia con to a defective low-density lipoprotein receptor, phenylketonuria due to phenylalanine hydroxylase deficiency, Paconi's anaemia, apolipoprotein crithine transcarbamylase deficiency, Faconi's anaemia, apolipoprotein
                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Histidine copolymer; enzyme deficiency; phenylalanine hydroxylase; ornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; purine nucleoside phosphorylase; alphal-antitrypsin; apolipoprotein E; chronic granulomatous disease; sickle cell; beta-thalassaemia; anaemia; familial hypercholesecrolaemia; low-density lipoprotein receptor; phenylketonuria; Faconi's anaemia; haemophilia; muscular dystrophy; cystic fibrosis; Parkinson's disease; retinitis pigmentosa; cytostatic; lysosomal storage disease; mycopolysaccharide type I disease; cardiant; diabetic retinopathy; human immunodeficiency virus disease; infection; cardiac disease; peripheral vascular disease; antisickling; arthritis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues
                                                                                                                           Gaps
                                                                                                                           ö
                                                                                  Length 21;
                                                                                                                           Indels
                                                                                                   Pred. No. 4.1e-09;
                                                                                65.6%; Score 120; DB 22;
100.0%; Pred. No. 4.1e-09;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               nootropic, haemostatic, virucide, gene therapy
                                                                                                                                                                                                                                                                                                             AAE06241 standard; peptide; 21 AA
                                                                                                                                                                                          2 KHKHKHKHKGKHKHKHK 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 3; Page 36; 64pp; English.
                                                                                                                                                                1 КНКНКНКНКСКНКНКНКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                      Histidine copolymer, peptide #4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0173576.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20-DEC-2000; 2000WO-US34603
                                                                                                                                                                                                                                                                                                                                                                                             25-SEP-2001 (first entry)
                                                                                                   Best Local Similarity 100.
Matches 19; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2001-425579/45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (MIXS/) MIXSON A J.
                                         21 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO200147496-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29-DEC-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      05-JUL-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mixson AJ;
                                                                                                                                                                                                                                                                                                                                                        AAE06241;
                                           Sequence
                                                                                  Query Match
K-HK.
                                                                                                                                                                                                                                                                      RESULT 7
```

```
Histidine copolymer; enzyme deficiency; phenylalanine hydroxylase; ornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; purine nucleoside phosphorylase; alphal-antitrypsin; apolipoprotein E; chronic granulomatous disease; alphal-antitrypsin; apolipoprotein E; familial hypercholesterolaemia; low-density lipoprotein receptor; phenylketonuria; Faconi's anaemia; haemophilia; muscular dystrophy; cystic fibrosis; Parkinson's disease; retinitis pigmentosa; cytostatic; lysosomal storage disease; mycopolysaccharide type I disease; cardiant; diabetic retinopathy; human immunodeficiency virus disease; infection; cardiac disease; peripheral vascular disease; anthitis;
E deficiency, haemophilia A and B due to factor VIII and IX deficiency, muscular dystrophy due to dystrophin, laminin-2, or sacroglycans mustations, cystic fibrosis due to CFTR mutations, Parkinson due to tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storag disease (i.e., mycopolysaccharide type 1, Hunter, Hurler and Gaucher), diabetic retinopathy, human immunodeficiency virus disease, virus infection, acquired anaemia, cardiac and peripheral vascular disease and arthritis. The present peptide sequence is a histidine copolymer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The invention relates to a pharmaceutical agent delivery composition comprising a transport polymer comprising a linear or branched peptide having about 10 to 300 amino acid residues, having about 5 to 100% histidine residues, and optionally having 1 to 95% non-histidine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                       Length 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                       65.6%; Score 120; DB 22; 100.0%; Pred. No. 4.1e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               nootropic; haemostatic; virucide; gene therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                             100.0%; Prec. ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1..11
/note= "Y component"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Histidine copolymer, Y-HK peptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAE06232 standard; peptide; 29 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example; Page 23; 64pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 КНКНКНКНКСКНКНКНК 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             20-DEC-2000; 2000WO-US34603.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0173576.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25-SEP-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity 100.
188 19; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2001-425579/45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A J.
                                                                                                                                                                                                                                                                                                                                   21 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200147496-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NIXS/) MIXSON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29-DEC-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      05-JUL-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mixson AJ;
                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
```

```
Haemophilus influenzae infection; BVH-NTHI1; otitis media; BVH-NTHI2; sinusitis; bronchitis; pneumonia; meningitis; bacteraemia; BVH-NTHI3; BVH-NTHI4; BVH-NTHI5; BVH-NTHI6; BVH-NTHI7; BVH-NTHI8; BVH-NTHI9; BVH-NTHI10; BVH-NTHI11; BVH-NTHI11; BVH-NTHI11; antiinflammatory; auditory;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Novel isolated Haemophilus influenzae polypeptides BVH-NTHI1-12, u
for inducing protective immune responses against H. influenzae in
animals and for treating otitis media, sinusitis, bronchitis and
                                                                                                                                                                                                              Score 117; DB 22; Length 29;
Pred. No. 1.4e-08;
4; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouellet C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Brodeur BR, Martin D,
                                                                                                                                                                                                                                                                                                                                                                             H influenzae BVH-NTHI2 protein SEQ ID NO: 4.
                                                                                                                                                                                                                                                      1 КНКНКНКНКСКНКНКНКНКНК 27
                                                                                                                                                                                                                                                                    RKKRRQRRRGKHKHKHKHKHKHKHK 29
residues, at least one pharmaceutical
                                                                                                                                                                                                                                                                                                                        AA017803 standard; Protein; 335 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 17; Fig 4; 58pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  02-OCT-2000; 2000US-236712P.
                                                                                                                                                                                                                    63.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                02-OCT-2001; 2001WO-CA01402.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Charbonneau A, Vayssier C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (SHIR-) SHIRE BIOCHEM INC.
                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                       18; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Haemophilus influenzae.
                                                                                                                                                                                                                                                                                                                                                                                                                                     antibacterial; vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hamel J, Couture F,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2002-435325/46.
                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAL46801.
                                                                                                                                                                                                  29 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200228889-A2.
                                                                                                                                                                                                                                                                                                                                                               05-AUG-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11-APR-2002.
                                                                                                                                                                                                                                                                                                                                            AAO17803;
                                                                                                                                                                                                    Seguence
                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                                                                                                                                             8$8888888888888888888888888888
                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                          ò
```

Gaps

ö

```
ö
The present invention provides the protein and coding sequences of Haemophilus influenzae BVH-NTHI1-12. The sequences can be used in the production of a vaccine to protect against, and in the diagnosis of, H. influenzae infection, which can lead to otitis media, sinusitis, bronchitis, pneumonia, meniatis and bacteraemia. The present sequence is a protein of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     be used
                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This is the amino acid sequence of the 46 kDa cell surface adhesin protein of Haemophilus influenzae. The adhesin binds specifically to phosphatidylethanolamine, gangliotriaosyloeramide and supplicternasyloeramide. It can be purified from H. influenzae cells or from transformed host cells. Adhesin polypeptides or nucleic acid molecules (see ANV36489) encoding them can also be use in claimed immunogenic compositions formulated as vaccines to protect a mammal (preferably a human) against diseases caused by bacterial pathogens that have the adhesin as a surface protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Haemophilus influenzae adhesin protein – useful in vaccines against
bacțerial infection e.g. H. influenzae or E. coli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62.3%; Score 114; DB 19; Length 337; 60.7%; Pred. No. 3.9e-07; ive. 3; Mismatches 8; Indels (
                                                                                                                               Length 335;
                                                                                                                                                              Indels
                                                                                                                                 Score 115; DB 23;
Pred. No. 2.9e-07;
3; Mismatches 7;
                                                                                                                                                                                                          129 HKHEHKHDHKHDHKHEHKHDHEH 155
                                                                                                                                                                                         2 нхнкнкнкскнкнкнкнкнкн 28
                                                                                                                                                                                                                                                                                                                                                                             Haemophilus influẹnza adhesin protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1..24
/label= Sig_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 6; Page 32; 48pp; English
                                                                                                                                                                3;
                                                                                                                                                                                                                                                                                           AAW61155 standard; Protein; 337
                                                                                                                                     63.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         96CA-2182046.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        96CA-2182046.
                                                                                                                                                                                                                                                                                                                                                                                                         Adhesin; vaccine; infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 62.3
Best Local Similarity 60.7
Matches 17; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (HSCR-) HSC RES & DEV LP.
                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                     Haemophilus influenzae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1998-333955/30.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                337 AA;
                                                                                                                                                    Local Similarity
ses 17; Conserv
                                                                                                              335 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAV36458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25-JUL-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        25-JUL-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Lingwood CA;
                                                                                                                                                                                                                                                                                                                                                    12-OCT-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CA2182046-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26-JAN-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                           AAW61155;
                                                                                                              Sequence
                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Peptide
                                                                                                                                                        Best Loc
Matches
                                                                                                                                                                                                                                                                    RESULT 10
                                                                                                                                                                                                                                                                               AAW61155
                                                                                                                                                                                                                                                                                                             8888888888
                                                                                                                                                                                                 ò
                                                                                                                                                                                                                         셤
```

ö

2 НКНКНКНКККНКНКНКНКНКНКНКНК 29

à

useful

Tremblay M;

Ź

```
129 HKHEHKADHKADHOHDHDHKHEHKIDHE 156
                                        Human ovarian cancer markers M605-M607
                    ABG96401 standard; Protein; 1043
                                  (first entry)
                                                                                                                                                                essing the stage cexpression level
                                                                                                                                                  2002-723277/78
                                                                                WO200271928-A2.
                                                                         Homo sapiens.
                                                                                                    14-MAR-2001;
                                  11-DEC-2002
                                                                                                                                     Monahan JE,
Meyers RE,
                                                                                       19-SEP-2002
                           ABG96401;
                                                                                                                                            Bast RC,
   용
```

patient is afflicted with overian cancer: The method involves comparing the expression level of a marker in a patient sample and the normal level of expression of the marker in a control non-ovarian cancer sample, where the marker is selected from 33 cancer markers described in the specification. The method of the invention is useful in diagnosing or characterising cancer, in detecting the presence of cancer as early as possible, and the recurrence of ovarian cancer. The method may also be particular use with patients having an enhanced risk of developing ovarian cancer (e.g. patients having a familial history of ovarian cancer?). The cancer markers may be used in the management and treatment of e.g. brain and central nervous system disorders (e.g. bacterial and viral meningitis Alzheimer's disease or Parkinson's disease), brain disorders (e.g. cerebral oedema, hydrocephalus or brain herniations), inflammations (e.g. bacterial or viral meningitis or encephalitis), central nervous system disorder; bacterial meningitis; viral meningitis; Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus; Barain herniation; inflammation; encephalitis; testicular disorder; nontuberculous granulomatous orchitis; connective tissue disorder; heart disorder; ischaemic heart disease; atherosclerosis; neoplasm; histological type; carcinogenic; ovarian cancer marker. ovarian cancer; marker; cancer; familial history; brain disorder; Assessing whether a patient is afflicted with ovarian cancer, useful in assessing the stage or progression of the disease, comprises comparing or progression of the disease, comprises comparing I of a cancer marker in a sample from a patient and testicular disorders (e.g. nontuberculous granulomatous orchitis), connective tissue disorders, or heart disorders (e.g. ischaemic heart disease or atherosclerosis). The compositions and methods may also be used in assessing the histological type of neoplasm associated with The present invention relates to a new method for assessing whether Hoersch S, Kamatkar S, Kovatis SG; andt PJ, Sen A, Vieby PO, Mills GB; Zhao X, Glatt K; Disclosure; Page 378-381; 481pp; English. Gannavarapu M, Hoere Town MP, Olandt PJ, Se N-PSDB; ABS76498, ABS76499, ABS76500 Lu K, Schmandt RE, (MILL-) MILLENNIUM PHARM INC 14-MAR-2001; 2001US-276026P. 10-AUG-2001; 2001US-311732P. 19-SEP-2001; 2001US-323580P. 26-SEP-2001; 2001US-324967P. 26-SEP-2001; 2001US-325102P. 26-SEP-2001; 2001US-325149P. 14-MAR-2002; 2002WO-US07826 from a non cancer patient -Morrisey MP,

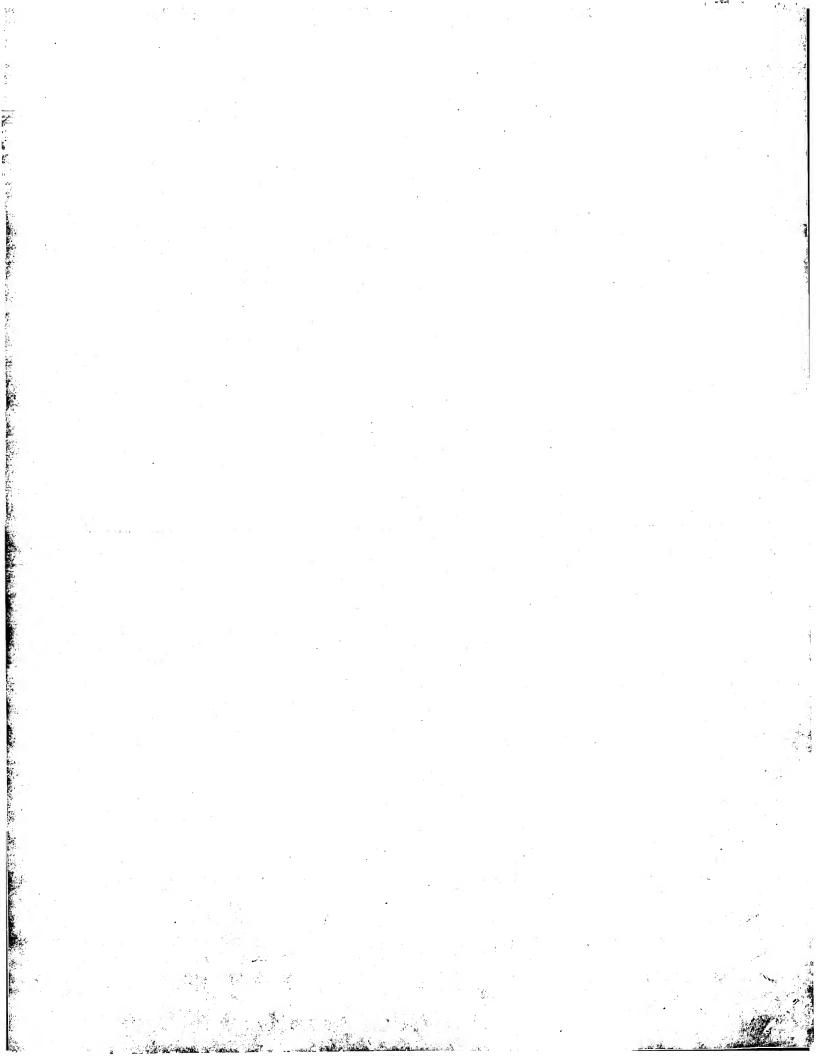
```
3
ovarian cancer, monitoring the progression of ovarian cancer, determining whether ovarian cancer has metastasized or is likely to metastasize, selecting a composition for inhibiting ovarian cancer, assessing the ovarian cancer optential of a compound, or inhibiting ovarian cancer or at risk of developing ovarian cancer represents one of the ovarian cancer markers described in the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sequence and an oligonucleotide comprising a sequence complementary to a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the
                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human, primer; detection; diagnosis; antisense therapy; gene therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Yamamoto J;
                                                                                                                                                                                 60.4%; Score 110.5; DB 23; Length 1043; 59.0%; Pred. No. 3.4e-06; ive 1; Mismatches 4; Indels 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Saito K,
Otsuki'
                                                                                                                                                                                                                                                                               1 KH----KHKHKHKHKHKHKHKHKG-----KHKHKHK 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 8; SEQ ID 12708; 2537pp + CD ROM; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hayashi K, S
A, Nagai K,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Human protein sequence SEQ ID NO:12708
                                                                                                                                                                                                                                                                                                                                                                                       AAB93454 standard; Protein; 271 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wakamatsu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Nishikawa T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27-AUG-1999; 99JP-0300253.
11-JAN-2000; 2000JP-0118776.
02-MAX-2000; 2000JP-0183767.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28-JUL-2000; 2000EP-0116126.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09-JUN-2000; 2000JP-0241899
                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                        23; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (HELI-) HELIX RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sugiyama T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2001-318749/34.
                                                                                                                                                   1043 AA;
                                                                                                                                                                                                       Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Isogai T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Full-length cDNAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             26-JUN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EP1074617-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      29-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              07-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4 T,
                                                                                                                                                                                                                                                                                                                                                                                                                             AAB93454;
                                                                                                                                                    Seguence
                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ishii
                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ota
                                                                                                                                                                                                                                                                                                                                                                    AAB93454
     88888888888
                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                            8
```

```
ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The present invention describes primer sets for synthesising 5602 full-length CDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprises a 3'-end sequence the oligonucleotide comprises a 1-east 15 nucleotides and the complementary to a polynucleotide comprises at 1-east 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and
            detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAS. The primers allow obtaining of the full-length cDNAS easily without any specialised methods. AAM10316c to AAM13628 and AAM13633 to AAM18742 represent human enthous any sequences; AAM3546 to AAM35893 represent human amino acid sequences; and AAM13629 to AAM13632 represent oligonucleotides, all of which are used in the exemplification
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human; primer; detection; diagnosis; antisense therapy; gene therapy.
                                                                                                                                                                                                                            Gaps
The primers are also useful for the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Yamamoto J;
                                                                                                                                                                                                                            1;
                                                                                                                                                                                           DB 22; Length 271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Saito K, Ya
, Otsuki T;
                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                            7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8; SEQ ID 12031; 2537pp + CD ROM; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Wakamatsu A, Nagai K,
                                                                                                                                                                                          Score 104.5; DB
Pred. No. 5e-06;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hayashi K,
                                                                                                                                                                                                                             1; Mismatches
                                                                                                                                                                                                                                                                                 220 HHHHHEHK-KKKKHKHKHKHKHKHDSK 246
                                                                                                                                                                                                                                                            2 НКНКНКНКККНКНКНКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Human protein sequence SEQ ID NO:12031.
                                                                                                                                                                                                                                                                                                                                                                                    ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sogai T, Nishikawa T,
Sugiyama T, Wakamatsu
                                                                                                                                                                                                                                                                                                                                                                                  AAB93135 standard; Protein; 504
       particularly full-length cDNAs.
                                                                                                                                                                                             57.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         29-JUL-1999; 99JP-0248036.
27-AUG-1999; 99JP-0300253.
11-JAN-2000; 2000JP-0118776.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28-JUL-2000; 2000EP-0116126.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           02-MAY-2000; 2000JP-0183767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09-JUN-2000; 2000JP-0241899
                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                               Local Similarity 67.9
                                                                                                                             of the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (HELI-) HELIX RES INST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2001-318749/34.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    full-length cDNAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Isogai T,
                                                                                                                                                             271 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EP1074617-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            07-FEB-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      26-JUN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   s,
                                                                                                                                                                                                                                                                                                                                                                                                                      AAB93135;
                                                                                                                                                                 Sequence
                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ishii
                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                       RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                 셤
        88888888888888
```

```
This represents a human CIFISO/hTAFIIISO (CT) polypeptide that is conversed in mitosis and cell cycle progression. The CT protein conversed in mitosis and cell cycle progression. The CT protein correctly conversed in the CT protein plays a role in establishing gene expression conversed to patterns necessary for progression through the cell cycle. CT is a necessary transcriptional regulator of cell cycle progression through the cell cycle. CT is a converse on the used for affecting mitosis or cell cycle progression and treating disorders associated with alterations in mitosis correctly correctly correctly correctly disorders associated with alterations in mitosis correctly correctly correctly benefits and progression. Decreased CT gene expression can be used to creat conditions characterised by high rates of mitosis, such as conversed 
        in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13631 to AAH18742 represent human cDNA sequences; AAB92446 to AAH13632 represent human amino acid sequences; and AAH13629 to AAH13632 represent human amino acid sequences; and AAH13629 to AAH13632 represent invention.
The primers are useful for synthesising polynucleotides,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CIF150/hTAFI1150; mitosis; cell cycle progression; neoplasia; human;
transcriptional regulator; benign proliferative disease; dysplastic;
                                                                                                                                                                                                                                                                                                                                                                   1,
                                                                                                                                                                                                                                                                                                                  Length 504;
                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                     DB 22;
                                                                                                                                                                                                                                                                                                                  Score 104.5; DB 22;
Pred. No. 9.4e-06;
1; Mismatches 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New isolated lCIF150/hTAFII150 polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                               HHHHHEHK-KKKKKHKHKHKHKHDSK 479
                                                                                                                                                                                                                                                                                                                                                                                                                            2 НХНКНКНКНКНКНКНКНКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY29083 standard, Protein; 1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 2; Page 78; 82pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Human CIF150/hTAFII150 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           transcriptional regulator; ben
hyperplastic disorder; anemia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US27665.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97US-0068956.
                                                                                                                                                                                                                                                                                                                          57.18;
                                                                                                                                                                                                                                                                                                                                                   67.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                              19; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI: 1999-444069/37.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                        Local Similarity
                                                                                                                                                                                                                                                                            504 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAX91483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO9933985-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30-DEC-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    27-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-DEC-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08-JUL-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Kaufmann J;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY29083;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  453
                                                                                                                                                                                                                                                                               Seguence
                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY29083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                88888888888888
                                                                                                                                                                                                                                                                                                                                                                                                                                    ਨੇ
```

```
Job time : 45.971 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ä
                                                                                                                       1;
either in vitro, such as in a cell culture, or in vivo. The products can also be used for the diagnosis, prognosis and predicting susceptibility to neoplasia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The present sequence is a human CIF150/hTAFI1150 protein which is required for a human cell to enter mitosis. In particular, CIF150/hTAFI1150 is an essential cofactor for TPIID-dependent transcription. CIF130 negatively requlates CIF150 thereby inhibiting mitosis or cell cycle progression. CIF150 has a CIF130-binding domain and is used to screen compounds that interfere with the binding of CIF130. CIF130 is used to diagnose and treat proliferative disorders including hyperplasias, neoplasias and dysplasias.
                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1; Gaps
                                                                                                                                                                                                                                                                                                                                                                                  Human, CIP130 protein, cell cycle regulator; mitosis; neoplasia;
CIF150/hTAFI1150; diagnosis; treatment; proliferative disorder;
hyperplasia; dysplasia; CIF150.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 104.5; DB 21; Length 1199;
Pred. No. 2.3e-05;
1; Mismatches 7; Indels 1;
                                                                                     Score 104.5; DB 20; Length 1199;
Pred. No. 2.3e-05;
1; Mismatches 7; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Novel regulatory protein, useful to diagnose and treat cell proliferation disorders, including cancer -
                                                                                                                                                                                                                                                                                                                                                         Human CIF150/hTAFII150 essential for mitosis.
                                                                                                                                                                       2 НКНКНКНКСКНКНКНКНКНКНКНКНК 29
                                                                                                                                                     2 НКНКНКНКСКНКНКНКНКНКНКНКНК 29
                                                                                                                                                                                                                                                             AAY44263 standard; Protein; 1199 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 20; Fig 3; 52pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 67.9%;
Matches 19; Conservative 1
                                                                                       Query Match 57.1%;
Best Local Similarity 67.9%;
Matches 19; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99WO-US13143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0089198.
98US-0111636.
                                                                                                                                                                                                                                                                                                                           22-MAR-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2000-116524/10.
N-PSDB; AAZ29605.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (CHIR ) CHIRON CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1199 AA;
                                                              Sequence 1199 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12-JUN-1998;
09-DEC-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16-DEC-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kaufmann J;
                                                                                                                                                                                                                                                                                             AAY44263;
 8X3333
                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
```

Search completed: January 20, 2004, 18:23:18



```
US-08-686-528A-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-686-528A-3
                                                                                              69.5
64.5
64.5
64.5
64.5
64.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 1, Appli
Sequence 4, Appli
Sequence 6, Appli
Sequence 64, Appl
Sequence 64, Appl
Sequence 64, Appl
Sequence 1524, Appl
Sequence 2356, A
Sequence 34, Appli
Sequence 34, Appli
Sequence 34, Appli
Sequence 25591, A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Appl
Appli
Appli
Appli
Appli
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                               January 20, 2004, 18:22:04; Search time 16.3913 Seconds (without alignments) 74.858 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 3,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 3,
Sequence 2,
Sequence 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 2,
Sequence 4,
Sequence 2,
Sequence 2,
Sequence 6,
Sequence 1,
Sequence 1,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 1
Sequence 1
Sequence 6
                                                                                                                                                                                                                                                                                                                                                                  Issued Patents AA:*

(cgn2_6/ptodata/1/iaa/5A_COMB.pep:*

(cgn2_6/ptodata/1/iaa/5B_COMB.pep:*

(cgn2_6/ptodata/1/iaa/6A_COMB.pep:*

(cgn2_6/ptodata/1/iaa/6B_COMB.pep:*

(cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*

(cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*

(cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
          GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-686-528A-3
US-09-456-287-3
US-09-456-287-3
US-09-456-287-3
US-09-456-287-2
US-09-312-25-4
US-09-312-25-4
US-09-312-25-4
US-09-517-605-6
US-09-517-605-6
US-09-612-126-1
US-09-612-126-1
US-09-612-126-1
US-09-612-126-1
US-09-612-126-4
US-09-612-126-4
US-09-612-126-4
US-09-612-126-4
US-09-612-126-4
US-09-612-126-4
US-09-612-126-4
US-09-612-126-4
US-09-613-126-4
US-09-613-126-4
US-09-613-126-4
US-09-613-126-4
US-09-613-126-4
US-09-613-126-4
US-09-613-126-4
US-09-613-126-4
US-09-613-126-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-252-991A-23560
US-09-602-565-34
US-09-252-991A-25291
                                                                                                                                   US-10-018-103A-4
183
1 XHKHKHKHKHKGKHKHKHKKHK 29
                                                                                                                                                                                                                                                     Fotal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                            328717 segs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUMMARIES
                                                                                                                                                                                                                                                                                                                Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                            OM protein - protein search, using sw model
                                                                                                                                                                                       BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                            Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query
Match Length
                                                                                                                                     Title:
Perfect score:
Sequence:
                                                                                                                                                                                        Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           114
104.5
104.5
104.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83.5
83.5
80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           79.5
79.5
79.5
78.5
                                                                                                                                                                                                                            Searched:
                                                                                                                                                                                                                                                                                                                                                                     Database
                                                                                     Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Result
No.
```

```
29 745 40.7 1213 1 UG-06 646 778-20 Sequence 20.0 Appl 29 775 39.6 1664 2 UG-09 5667 718-20 Sequence 20.0 Appl 20 772 59.8 6 1664 2 UG-09 569 662-2 Sequence 20.0 Appl 20 772 59.8 6 1664 2 UG-09 599 662-2 Sequence 20.0 Appl 20 772 59.8 6 1664 2 UG-09 599 662-2 Sequence 20.0 Appl 20 772 59.8 6 1664 2 UG-09 599 662-2 Sequence 20.0 Appl 20 772 59.8 6 1664 2 UG-09 599 662-2 Sequence 20.0 Appl 20 772 59.8 6 1664 2 UG-09 599 662-2 Sequence 20.0 Appl 20 772 59.8 6 1664 2 UG-09 599 662-2 Sequence 20.0 Appl 20 772 59.8 6 1664 2 UG-09 70 1664 2 UG
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
62.3%; Score 114; DB 3; Length 337;
Best Local Similarity 60.7%; Pred. No. 5.1e-07;
Matches 17; Conservative 3; Mismatches 8; Indels
                                                                                                                                                                                                                                                                                                                                                                     Length 337;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SUCURNICES:

NUMBER OF SUCURNICES:

ADORESPONDENCE ADDRESS:

ADDRESSE:
Burns, Doane, Swecker & Mathis, L.L.P.

STREET:
STREET:
STREET:
STATE:
STAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 4
US-09-456-287-2
US-09-456-2. Application US/09456287
; Sequence 2, Application US/09456287
; Patent No. 6218147
; GENERAL INFORMATION:
APPLICANT: LINGWOOD, Clifford A.
ITLE OF INVENTION: HARMOPHILUS ADHESIN PROTEIN
                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 60.7%; Pred. No. 5.1e-07;
Matches 17; Conservative 3; Mismatches 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             129 HKHEHKHDHKHDHDHDHKHEHKHDHE 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2 НКНКНКНКСКНКНКНКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 НКНКНКНКНКНКНКНКНКНКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    032609-001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/686,528
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 2, Application US/09208742
Patent No. 6174679
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATTORNEY/AGENT INFORMATION:
NAME: Rea, Teresa Stanek
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 0326
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              337 amino acids
                                                                                           337 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Kaufmann, Joerg
INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; MOLECULE TYPE: protein US-09-456-287-2
                                                                                                                                                                     single
                                                                                                                                                                                                                 TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY: linear
                                                                                                                              TYPE: amino acid
STRANDEDNESS: sin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 5
US-09-208-742-2
                                                                                                                                                                                                                                                                                                        US-08-686-528A-2
                                                                                               LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ठ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 114; DB 3; Length 313;
Pred. No. 4.7e-07;
3; Mismatches 8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-886-28A-2

S-GY-86B-22BA-2

S-GY-86B-22BA-2

S-GY-86C-2

SEQUENCE 2, Application US/0868652BA

Patent No. 6054134

GENERAL INFORMATION:

TITLE OF INVENTION: HAEMOPHILUS ADHESIN PROTEIN
NUMBER OF SEQUENCES:

ADDRESSE: Burns, Doane, Swecker & Mathis, L.L.P.

STREET: 1737 King Street, Suite 500

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

SIAPE: 2314-2756

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATE:

APPLICATION NUMBER: US/08/686,528A

FILING DATE: 26-ULL.1996

CLASSIFICATION NATE:

SALUL.1996

CLASSIFICATION NATE:

CLASSIFICATION SALUL.1906
                          GENERAL INFORMATION:
APPLICANT: LINGWOOD, Clifford A.
APPLICANT: LINGWOOD, Clifford A.
TITLE OF INVENTIONS: 4
CORRESPONDENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSE: Burns, Doane, Swecker & Mathis, L.L.P.
STREET: 1737 King Street, Suite 500
CITY: Alexandria
CTITY: Alexandria
CTITY: Alexandria
STATE: Virginia
COMPTRY: United States
COMPUTER: Flopy disk
COMPUTER: Flopy disk
COMPUTER: Flopy disk
COMPUTER: PLADES FORM:
MEDIUM TYPE: Flopy disk
COMPUTER: PLADES FORM:
MEDIUM TYPE: PLOPY DISK
COMPUTER: PLADES FORM:
MEDIUM TYPE: PLOPY DISK
COMPUTER: PLADES FORM:
MEDIUM TYPE: PLADES FORM:
MEDIUM 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               105 HXHEHKHDHKHDHDHDHDHKHEHKHDHE 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
ATTORNEY/ACENT INFORMATION:
NAME: Rea, Teresa Stanek
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 032609-001
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 313 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 НҮНКНҮКНКӨКНКНҮНКНКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CLASSIFLATION:
ATTORNEY/AGENT INFORMATION:
NAME: Rea, Teresa Stanek
REGISTRATION NUMBER: 30,427
REFERENCE/DOCKET NUMBER: 032609-001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/686,528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/456,287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 62.3%;
Best Local Similarity 60.7%;
Matches 17; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS: single TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-686-528A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-456-287-3
```

ઠ 셤

요 ò

```
RESULT 9

US-09-476-482-2

i Sequence 2, Application US/09476482

patent No. 628456

GENERAL INFORMATION:
APPLICANT: Jones, Katherine A.
APPLICANT: Jones, Matherine A.
APPLICANT: Application US/09476482

APPLICANT: Ping
APPLICANT: Ping
APPLICANT: Pang, Shi.-Min
TITLE OF INVENTION: A TRANSCRIPTIONAL COACTIVATOR THAT
TITLE OF INVENTION: RNA, METHODS FOR MODULATING TAT TRANSACTIVATION, AND USES
TITLE OF INVENTION: TRANS. US/09/476,482

CURRENT APPLICATION NUMBER: US/09/126,980

EARLIER FILING DATE: 1999-12-30

EARLIER FILING DATE: 1998-07-30
                                                                                                                                                    Sequence 2, Application US/09126980
; Sequence 2, Application US/09126980
; Patent No. 6270956
; GENERAL INFORMATION:
    APPLICANT: Jones, Katherine
; APPLICANT: Garber, Mitchell
; APPLICANT: Fands, Shi-Min
; TITLE OF INVENTION: INTERACTS WITH TAT PROTEIN AND REGULATES ITS
; TITLE OF INVENTION: INTERACTS WITH TAT PROTEIN AND REGULATING TAT
; TITLE OF INVENTION: INTERACTS WITH TAT PROTEIN AND REGULATING TAT
; TITLE OF INVENTION: TRANSACTIVATION, AND USES THEREFOR
; FILE REFERENCE: SALK2231
; CURRENT APPLICATION NUMBER: US/09/126,980
; EARLIER APPLICATION NUMBER: 60/069,341
; GURBERT FILING DATE: 1999-07-30
; EARLIER PILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH 726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 726;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 84; DB 3; Length /20
Pred. No. 0.0039;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 84; DB 3;
Pred. No. 0.0039;
                                   1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   505 EHKEKHKTHPSNHHHHNH---HSHKHSH 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        505 EHKEKHKTHPSNHHHHNH---HSHKHSH 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 КНКНКНК-НКСКНКНКНКНКНКНКН 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 КНКНКНК-НКСКНКНКНКНКНКНКНКНКН 28
       2 НКНКНКНКСКНКНКНКНКНКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 51.7%;
Matches 15; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 51.7%;
Matches 15; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: Homo sapiens
US-09-126-980-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; ORGANISM: Homo sapiens
US-09-476-482-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠ
         ઠે
                                                                                                                                                                                                                                                                                                                                                       ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 1199;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 1199;
                                                                                                                                                                                                                                                                                                         Length 1199;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GENERAL INFORMATION:
APPLICANT: KAUFMAIN, JOSTY
TITLE OF INVENTION: CIF130 INHIBITS CELL CYCLE PROGRESSION
FILE REFERENCE: 200130.456 / 1513.003
CURRENT APPLICATION NUMBER: US/09/709,979
CURRENT FILING DATE: 2000-11.09
PRIOR APPLICATION NUMBER: US 09/332,295
PRIOR FILING DATE: 1999-06-11
TITLE OF INVENTION: CIF150/hTAFI1150 is Necessary for Cell TITLE OP INVENTION: Cycle Progression PILE REFERENCE: 1453.002
CURRENT APPLICATION NUMBER: US/09/208,742
CURRENT FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Kaufmann, Joerg
TITLE OF INVENTION: CIF130 INHIBITS CELL CYCLE PROGRESSION
FILE REFERENCE: 200130,456 / 1513,003
CURRENT APPLICATION NUMBER: US/09/332,295
CURRENT FILING DATE: 1999-06-11
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FRRESEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                       7; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     57.1%; Score 104.5; DB 4;
67.9%; Pred. No. 2.1e-05;
tive 1; Mismatches 7;
                                                                                                                                                                                                                                                                                                         DB 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 57.1%; Score 104.5; DB 4; Best Local Similarity 67.9%; Pred. No. 2.1e-05; Matches 19; Conservative 1; Mismatches 7;
                                                                                                                                                                                                                                                                                                       Score 104.5; DB 3
Pred. No. 2.1e-05;
1; Mismatches 7
                                                                                                                                                                                                                                                                                                                                                                                                                               1148 HHHHHHK-KKKKKHKHKHKHKHKHKSK 1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 НКНКНКНКККНКНКНКНКНКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 НКНКНКНКСКНКНКНКНКНКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 4, Application US/09332295
Patent No. 6303372
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 4, Application US/0970979
Patent No. 6423822
                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 67.9%;
Matches 19; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 57.1
Best Local Similarity 67.9
Matches 19; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT
ORGANISM: Homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; ORGANISM: Homo sapien
US-09-332-295-4
                                                                                                                                                                                                           TYPE: PRT
ORGANISM: human
                                                                                                                                                                                        1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQ ID NO 4
LENGTH: 1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1199
                                                                                                                                                                                                                                                         US-09-208-742-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -09-332-295-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-709-979-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQ ID NO 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
```

ઠે g

```
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
Sequence 6, Application US/09517605
Sequence 6, Application US/09517605
Sequence 6, Application US/09517605
Barent No. 6331567
GENERAL INFORMATION:
APPLICAMT: Littuman, Dan R.
APPLICAMT: Kwon, Douglas S.
APPLICAMT: Wow, Douglas S.
APPLICAMT: Wow, Douglas S.
APPLICAMT: Geijtembeck, Tneo
TITLE OF INVENTION: CELLS
FILE REFERENCE: 1049-1-017
CURRENT APPLICATION NUMBER: US/09/517,605
CURRENT PILING DATE: 2000-03-02
NUMBER OF SEQ ID NOS: 17
                                                                                                                                                                                                                                                                                                                                                                                                                         7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----KGKH------KHKHKGKH------KHKHKH 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22 HKHKDKHKDREHRHKEHKEKDREKSKHSNSEHKDSEKKHKEKEKTKHKDGSSEKHKDKH 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              33;
                                                                                                                                                                                                                                                                                                                                                                                                                             4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 4; Length 300;
                                                                                                                                                                                                                                                                                                                                                                                    Score 84; DB 4; Length 726; Pred. No. 0.0039;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Hwang, Jaulang
APPLICANT: Hwang, Jaulang
APPLICANT: Hwang, Jaulang
APPLICANT: Hul, cho-Fat
APPLICANT: Chen, Taong-Yuaeh
APLICANTON: TOOPOISOMBRASE 1-MEDIATED DNA DELIVERY
FILE REPERBUCE: 089191/024001
CURRENT APPLICATION NUMBER: US/09/395,689
CURRENT FILING DATE: 1999-09-13
NUMBER OF SEQ ID NOS: 6
SOFTWARE FESTSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                             9; Indela
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                             1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      505 ЕНКЕКНКТНРЅИНННИН---НЅНКНЅН 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 83.5;
Pred. No. 0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 КНКНКНК-НКСКНКНКНКНКНКНКН 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: WAGATSUMA, Masako
APPLICANT: KURITA, No. 58495031ko
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 2, Application US/08663112; Patent No. 5849503; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Application US/09395689
Patent No. 6387684
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM: Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 36.1%;
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 51.7%;
Matches 15; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 HKHKHKH------
                                                                                                                                                                                                                                                                                                                    ; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-517-605-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29 K 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82 K 82
                                                                                                                                                                                                                                                                 SEQ ID NO 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQ ID NO 1
LENGTH: 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-663-112-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-395-689-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-395-689-1
JS-09-517-605-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

```
Sequence 10, Application US/09612126

Sequence 10, Application US/09612126

Sequence No. 6284710N:

GENERAL INFORMATION:

TITLE OF INVENTION:

TITLE OF INVENTION:

TITLE OF INVENTION:

TITLE OF INVENTION:

FILE REFERENCE: 6056-258 CT1

CURRENT FILING DATE: 105/09/612,126

CURRENT FILING DATE: 2000-07-07

PRIOR PAPLICATION NUMBER: 60/107,844

PRIOR APPLICATION NUMBER: 60/107,844

PRIOR FILING DATE: 1999-11-10

PRIOR FILING DATE: 1999-11-09

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 10

LENGTH: 94

THENCH APPLICATION NUMBER: 1201-109

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 10

LENGTH: 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22 HKHKDKHKDREHRHKEHKKEKDREKSKHSNSEHKDSEKKHKEKEKTKHKDGSSEKHKDKH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OTHER INFORMATION: Description of Artificial Sequence: Human high OTHER INFORMATION: molecular weight kininogen light chain amino acids OTHER INFORMATION: Lys(420) through Ser(513)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 765;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6; Indels
TITLE OF INVENTION: MUTANT PROTEINS OF HUMAN DNA TITLE OF INVENTION: TOPOISOMERASE I NUMBER OF SEQUENCES: 7 CORRESPONDENCE ADDRESS: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & ADDRESSEE: Dunner L.L.P. STREET: 1300 I Street, N.W.
                                                                                                                                                                                                                                                                                                                                   COMPUTER: TIEN PC compatible
COMPUTER: TIEN PC compatible
COMPUTER: TIEN PC compatible
COMPUTER: TIEN PC compatible
COMPUTER: Patentin Release #1.0, Version #1.30
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/663,112
FILING DATE: 26-NOV-1996
CLASSIFICATION: 435
ATTONNEY/AGENT INFORMATION:
NAME: Enhandi, Carolyn P.
REGISTRATION NUMBER: 32,220
REFERENCE/POCKET NUMBER: 06609.1488-00000
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  th 45.6%; Score 83.5; DB 2; Similarity 36.1%; Pred, No. 0.0047; 22; Conservative 0; Mismatches 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: PRT ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2 HKHKHKH------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : 765 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             202-408-4400
                                                                                                                                                                                                                                                           COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE: protein
                                                                                                                                                                                                     CITY: Washington STATE: D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              29 K 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82 K 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-663-112-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
```

81

```
g
                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
US-09-612-126-8

i Sequence 8, Application US/09612126

sequence 8, Application US/09612126

general INFORMATION:
f APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF
TITLE OF INVENTION: MOLECULAR WEIGHT KININGEN BOWAN 5

TITLE OF INVENTION: MOLECULAR WEIGHT KININGEN BOWAN 5

FILE REFERENCE: 6056-258 CT1
CURRENT FILING DATE: 2000-07-07

PRIOR APPLICATION NUMBER: 60/107,844

PRIOR APPLICATION NUMBER: 60/107,844

PRIOR FILING DATE: 1998-11-10

PRIOR FILING DATE: 1998-11-09

NUMBER OF SEQ ID NOS: 12

SOFTWARE: PATENTING DATE: 1999-11-09

LENGTH: 186
                                                                                                                                                                                                                                                                                                                     US-09-612-126-11

Sequence 11, Application US/09612126

Sequence 11, Application US/09612126

Sequence 11, Application US/09612126

Sequence 11, Application US/09612126

SENERAL INFORMATION:

APPLICANT: TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF

TITLE OF INVENTION: MOLECULAR WEIGHT KININGEN DOWAN 5

TITLE OF INVENTION: MOLECULAR WEIGHT KININGEN DOWAN 5

FILE REFERENCE: 6056-258 CT1

CURRENT APPLICATION NUMBER: US/09/612,126

SETOR FILING DATE: 1998-11-10

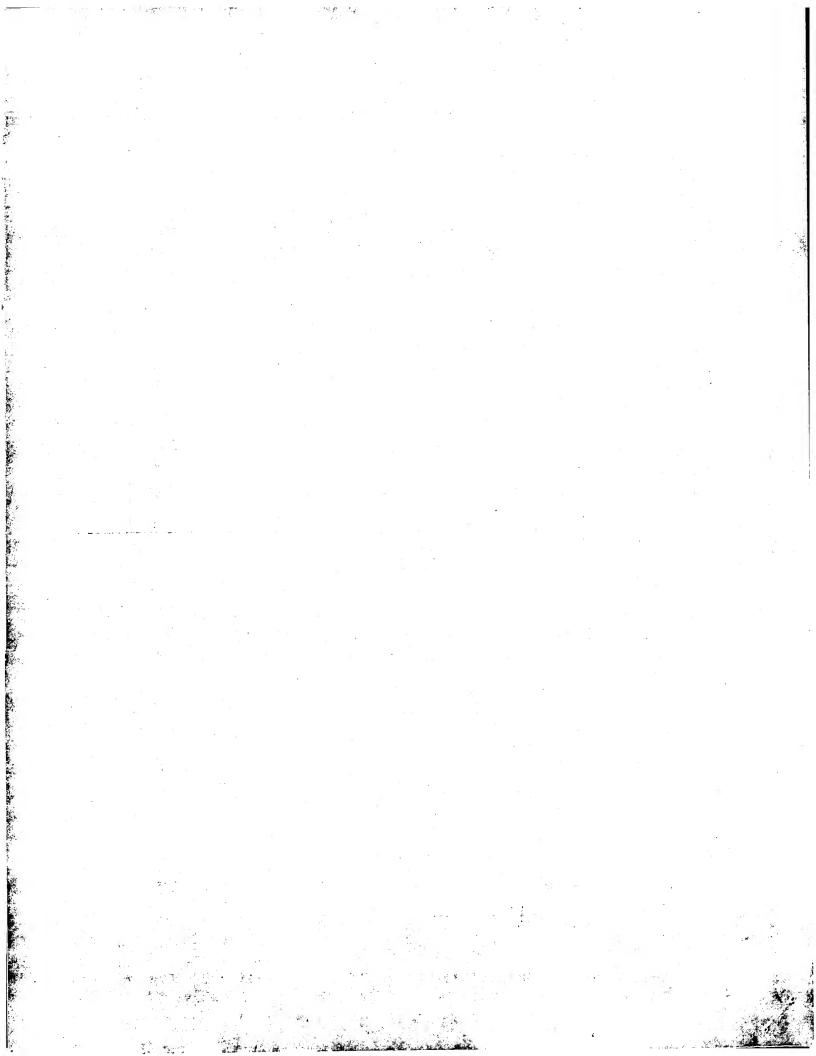
PRIOR PILING DATE: 1998-11-10

PRIOR PILING DATE: 1999-11-09

NUMBER OF SEQ ID NOS: 12

NUMBER OF SEQ ID NOS: 12
                                                                                                            7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OTHER INFORMATION: Description of Artificial Sequence: Human high OTHER INFORMATION: molecular weight kininogen light chain amino acids OTHER INFORMATION: His(441) through Ser(626)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OTHER INFORMATION: Description of Artificial Sequence: Human high OTHER INFORMATION: molecular weight kininogen light chain amino acids OTHER INFORMATION: Glu(448) through Ser(626)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4
                                                                                                               4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 43.7%; Score 80; DB 3; Length 179; Best Local Similarity 58.6%; Pred. No. 0.0034; Matches 17; Conservative 2; Mismatches 6; Indels
                                                   Length 94;
                                                                                                            6; Indels
                                                   Query Match 43.7%; Score 80; DB 3; Best Local Similarity 58.6%; Pred. No. 0.0019; Matches 17; Conservative 2; Mismatches 6
                                                                                                                                                                                               58 HKHKHGHGKGKKNGKKNGKTEH 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    30 HKHKHGHGHGKHKNKGKKNGKHNGWKTEH 58
                                                                                                                                                                2 НКНКНКН-КСКНКНКНКНКНКСКН---КНКН 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2 НКНКНКН-КСКНКНКНКНКККН---КНКН 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM: Artificial Sequence
US-09-612-126-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-612-126-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQ ID NO 11
LENGTH: 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-612-126-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: PRT
                                                                                                                                                                   ઠે
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

```
Query Match
Best Local Similarity 58.6%; Pred. No. 0.0035;
Matches 17; Conservative 2; Mismatches 6; Indels 4; Gaps 2;
Qy 2 HKHKHKHKKKHH---KHKH 26
Db 37 HKHKHGHGKHKNKKKKHOWKTEH 65
Search completed: January 20, 2004, 18:27:11
Job time: 16.3913 secs
```



```
(without alignments)
180.887 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                              January 20, 2004, 18:23:25; Search time 32.7826 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CGRIZ_6/PLOCATA 1/Pubpaa/USO7 PUBCOMB.pep:*

(CGRIZ_6/PLOCATA 1/Pubpaa/USO7 NEW PUB.pep:*

(CGRIZ_6/PLOCATA 1/Pubpaa/USO6 NEW PUB.pep:*

(CGRIZ_6/PLOCATA 1/Pubpaa/USO6 PUBCOMB.pep:*

(CGRIZ_6/PLOCATA 1/Pubpaa/USO6 PUBCOMB.pep:*

(CGRIZ_6/PLOCATA 1/Pubpaa/USO8 NEW PUB.pep:*

(CGRIZ_6/PLOCATA 1/Pubpaa/USO8 NEW PUB.pep:*

(CGRIZ_6/PLOCATA 1/Pubpaa/USO8 PUBCOMB.pep:*

(CGRIZ_6/PLOCATA 1/Pubpaa/USO8 PUBCOMB.pep:*

(CGRIZ_6/PLOCATA 1/Pubpaa/USO8 PUBCOMB.pep:*

(CGRIZ_6/PLOCATA 1/Pubpaa/USO8 PUBCOMB.pep:*

(CGRIZ_6/PLOCATA 1/Pubpaa/USO8 NEW PUB.pep:*

(CGRIZ_6/PLOCATA 1/Pubpaa/USO8 PUBCOMB.pep:*

(CGRIZ_6/PLOCATA 1/Pubpaa/USO8 NEW PUB.pep:*

(CGRIZ_6/PLOCATA 1/Pubpaa/USO8 NEW PUB.pep:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    prodata/1/pubpaa/0809B PUBCOMB.pep://prodata/1/pubpaa/0809B PUBCOMB.pep://prodata/1/pubpaa/0809B PUBCOMB.pep://prodata/1/pubpaa/0810A PUBCOMB.pep://prodata/1/pubpaa/0810A PUBCOMB.pep://prodata/1/pubpaa/0810A PUBCOMB.pep://prodata/1/pubpaa/0810B PUBCOMB.pep://pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB.pubcomB
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           183
1 КНКНКНКНКСКНКНКНКНККККНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Fotal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              762491 segs, 204481190 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Published Applications AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                     OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-10-018-103A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Searched:
                                                                                                                                                                                                                                                                                                                                                                                                            Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Result
```

Appli Appli Appl Appli Sequence 1406, Ap Sequence 4, Sequence 3, Sequence 3, Sequence 25 Sequence 4, Sequence 9 Sequence 1 Sequence 9 Sequence : Sequence Sequence Sequence Sequence Sequence US-10-131-909A-4
US-10-018-103A-4
US-10-131-909A-3
US-10-131-909A-3
US-10-131-909A-16
US-10-1018-103A-9
US-10-018-103A-9
US-10-018-103A-7
US-10-11-133-7
US-10-211-133-7
US-10-211-133-7
US-10-213-133-7
US-10-1338-279-4
US-10-138-279-4
US-10-138-279-4
US-10-138-279-4
US-10-138-279-4 110.5 110.5 104.5

16	94	51.4	15	77	US-10-131-909A-2	Sequence 2, Appli
17	94	51.4	12	15	US-10-018-103A-2	Sequence 2, Appli
18	93	50.8	337	15	US-10-270-333-96	Sequence 96, Appl
13	91	49.7	574	15	US-10-156-761-14106	Sequence 14106, A
20	90	49.2	19	12	US-10-131-909A-13	Sequence 13, Appl
21	90	49.2	19	15	US-10-018-103A-13	Sequence 13, Appl
22	98	47.0	106	15	US-10-106-698-6339	Sequence 6339, Ap
23	98	47.0	885	12	US-10-108-260A-3459	Sequence 3459, Ap
24	85	46.4	378	12	US-10-029-386-33892	Sequence 33892, A
25	84	45.9	78	12	US-10-195-730-186	Sequence 186, App
56	84	45.9	82	σ	US-09-864-761-33313	Sequence 33313, A
27	84	45.9	726	11	US-09-932-257A-19	Sequence 19, Appl
28	84	45.9	1257	12	US-10-369-493-6761	Sequence 6761, Ap
53	83.5	45.6	765	12	US-09-882-274-2	Sequence 2, Appli
30	82	44.8	4	σ	US-09-864-761-37882	Sequence 37882, A
31	81.5	44.5	20	12	US-10-131-909A-5	Sequence 5, Appli
32	81.5	44.5	20	15	US-10-018-103A-5	Sequence 5, Appli
33	81	44.3	13	12	US-10-131-909A-1	Sequence 1, Appli
34	81	44.3	13	15	US-10-018-103A-1	Sequence 1, Appli
35	81	44.3	444	12	US-10-360-849A-30	Sequence 30, Appl
36	80	43.7	148	σ	US-09-908-711-91	Sequence 91, Appl
37	80	43.7	148	=	US-09-764-891-3203	Sequence 3203, Ap
38	79.5	43.4	20	12	US-10-131-909A-6	Seguence 6, Appli
39	79.5	43.4	20	15	US-10-018-103A-6	
40	79	43.2	39	12	US-10-105-232-393	393,
41	79	43.2	39	12	US-10-189-437-380	380,
42	79	43.2	90	12	US-10-315-515-39	39, 4
43	79	43.2	90	12	US-10-315-515-44	Sequence 44, Appl
44	79	43.2	93	12	US-10-315-515-46	46,
45	79	43.2	95	12	US-10-315-515-35	35,

ALIGNMENTS

JS-10-131-909A-4

```
Gaps
                                                             JAPALICANT: Mixson, A. James
TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
FILE REFERENCE: 56274
FILE REFERENCE: 56274
CURRENT APPLICATION NUMBER: US/10/131,909A
CURRENT FILING DATE: 2001-11-05
FRIOR APPLICATION NUMBER: US 10/018103
FRIOR PILING DATE: 2001-11-05
FRIOR PILING DATE: 2000-12-20
FRIOR APPLICATION NUMBER: US 60/173576
FRIOR APPLICATION NUMBER: US 60/173576
FRIOR APPLICATION NUMBER: US 60/173576
FRIOR FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin Version 3.1
LENGTH: 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Match 100.0%; Score 183; DB 12; Local Similarity 100.0%; Pred. No. 1.4e-13; tes 29; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 КНКНКНКНКСКНКНКНКНКСКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OTHER INFORMATION: Synthetic Peptide
Sequence 4, Application US/10131909A; Publication No. US20030165567A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-10-131-909A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠ
```

ö

; Sequence 4, Application US/10018103A RESULT 2 US-10-018-103A-4

```
US-10-131-909A-9
                                                                                                                                                                                                                                                                                             FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-10-131-909A-3

| Sequence 3, Application US/10131909A
| Publication No. US20030165567A1
| GENERAL INFORMATION:
| APPLICANT: Mixson, A. James
| TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same | TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same | TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same | FILE REFERENCE: 5627*6
| CURRENT FILING DATE: 2001-11-05 | PRIOR FILING DATE: 2001-11-05 |
| PRIOR FILING DATE: 2001-11-05 |
| PRIOR FILING DATE: 2000-12-20 |
| PRIOR FILING DATE: 2000-12-20 |
| PRIOR FILING DATE: 1999-12-29 |
| NUMBER OF SEQ ID NOS: 17 |
| SEQ ID NO 3 |
| LENGTH: 19
Publication No. US20030045465Al
GENERAL INFORMATION:
APPLICANT: Mixson, A. James
TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
FILE REFERENCE: 5627*5
CURRENT APPLICATION NUMBER: US/10/018,103A
CURRENT APPLICATION NUMBER: US 60/173576
PRIOR APPLICATION NUMBER: PCT/US00/34603
PRIOR FILING DATE: 1999-12-29
PRIOR FILING DATE: 2000-12-20
PRIOR FILING DATE: 2000-12-20
NUMBER OF SEQ ID NOS: 16
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Version 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 3. Application US/10018103A
; Bublication No. US20030045465A1
; GENERAL INFORMATION:
; APPLICANT: Mixson, A. James
; TITLE OP INVENTION: Hatidine Copolymer and Methods For Using Same
; FILE REFERENCE: 5627*5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 100.0%; Pred. No. 5.2e-07;
Matches 19; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 100.0%; Pred. No. 1.4e-13;
Matches 29; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 КНКНКНКНКСКНКНКНКНКНКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 KHKHKHKHKGKRKHKHKHKHKHK 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; OTHER INFORMATION: Synthetic Peptide US-10-131-909A-3
                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-018-103A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 КНИНКНКНККИКНИНКНК 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 KHKHKHKHKGKHKHKHKHK 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-10-018-103A-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
```

```
Sequence 9, Application US/10131909A

Sequence 9, Application US/10131909A

Publication No. US20030165567A1

GENERAL INFORMATION:

APPLICANT: Mixson, A. James

TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same

TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same

TITLE OF INVENTION: WIMBER: US/10/131,909A

CURRENT APPLICATION NUMBER: US 10/018103

PRIOR APPLICATION NUMBER: PCT/US00/34603

PRIOR PILING DATE: 2001-11-05

PRIOR FILING DATE: 2001-12-29

PRIOR FILING DATE: 1999-12-29

NUMBER OF SEQ ID NOS: 17

SOFTWARE: PatentIn version 3.1

LENGTH: 21

TANDRE TO THE TO 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 16, Application US/10131909A
Publication No. US20030165567A1
GENERAL INFORMATION:
GENERAL INFORMATION:
TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same;
TITLE REPRENCE: 5527*6
CURRENT APPLICATION NUMBER: US/10/131,909A
CURRENT PILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: US 10/018103
PRIOR APPLICATION NUMBER: PCT/US00/34603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
65.6%; Score 120; DB 12;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 19; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 65.6%; Score 120; DB 15; Best Local Similarity 100.0%; Pred. No. 5.2e-07; Matches 19; Conservative 0; Mismatches 0;
CURRENT APPLICATION NUMBER: US/10/018,103A;
CURRENT FILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: US 60/173576
PRIOR FILING DATE: 1999-12-29
PRIOR FILING DATE: 1999-12-29
PRIOR FILING DATE: 2000-12-20
NUMBER OF SEQ ID NOS: 16
SEQ ID NO 3
SEQ ID NO 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OTHER INFORMATION: Synthetic Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OTHER INFORMATION: Synthetic Peptide US-10-018-103A-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2 XHKHKHKGKHKHKHKHK 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 КНКНКНКНКСКНКНКНКНК 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 КНКНКНКНКСКНКНКНК 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 KHKHKHKKKKKKKKKKKK 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: PRT ORGANISM: Artificial Sequence
```

```
US-10-131-909A-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-10-018-103A-7
                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 9, Application US/10018103A
; Sequence 9, Application US/10018103A
; Publication No. US20030045465A1
; Fublication No. US20030045465A1
; APPLICANT: Mixeon, A. James
; TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
; FILE REFERENCE: 5627*5
; CURRENT APPLICATION NUMBER: US/10/018,103A
; CURRENT FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: PCT/USO0/34603
; PRIOR APPLICATION NUMBER: PCT/USO0/34603
; PRIOR APPLICATION NUMBER: PCT/USO0/34603
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 16, Application US/10018103A
; Publication No. US20030045465A1
; GENERAL INFORMATION:
; APPLICANT: Mixson, A. James
; TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
; TITLE OF INVENTION: WUNBER: US/10/018,103A
; CURRENT APPLICATION WUNBER: US/10/018,103A
; CURRENT FILING DATE: 2001-11-05
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION WUNBER: PCT/US00/34603
; PRIOR APPLICATION WUNBER: PCT/US00/34603
; RIOR APPLICATION WUNBER: PCT/US00/34603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
65.6%; Score 120; DB 12; Length 21;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 19; Conservative 0; Mismatches 0;
                                      60/173576
                                                                                                                                                                                                                                                                                                                                           ; OTHER INFORMATION: Synthetic Peptide US-10-131-909A-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; OTHER INFORMATION: Synthetic Peptide US-10-018-103A-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 KHKHKHKHKGKHKHKHKHK 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2 KHKHKHKHKGKHKHKHK 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 КНКНКНКНКСКНКНКНК 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 КНКНКНКНКСКНКНКНК 19
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: US 60/
PRIOR FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin version 3.1
SEQ ID NO 16
                                                                                                                                                                                                                                                                               ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -10-018-103A-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -10-018-103A-9
                                                                                                                                                                                                                                                                                                            FEATURE:
                                                                                                                                                                                                            LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠ
```

```
ö
                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 7, Application US/10131909A
; Bachardon No. US20030165567A1
; GRENERAL INFORMATION:
    APPLICANTION:
    TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
    TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
; TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
; TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
; TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
; FILE REPRENCE: 562746
; CURRENT FILING DATE: 2001-11-05
; PRIOR FILING DATE: 2001-11-05
; PRIOR FILING DATE: 2001-12-20
; PRIOR PILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 7
; LENGTH: 29
                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Mixson, A. James
FILE REPERINCE: 5627*5
CURRENT APPLICATION WHERE: US/10/018,103A
CURRENT PILING DATE: 2001-11-05
PRIOR PAPLICATION NUMBER: US 60/173576
PRIOR PAPLICATION NUMBER: US 60/173576
PRIOR APPLICATION NUMBER: US 60/173576
PRIOR APPLICATION NUMBER: US 60/173576
PRIOR PELING DATE: 2000-12-20
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin version 3.1
SEQ ID NO 7
LENGTH: 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                           ő
                                                                                                                                                        Length 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 29;
                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                      Query Match
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 19; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
63.9%; Score 117; DB 12;
Best Local Similarity 66.7%; Pred. No. 1.6e-06;
Matches 18; Conservative 4; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 КНКНКНКНКККНКНКНКНКНКНК 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 RKCRRORRGKHKHKHKHKHKHKK 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; PEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-018-103A-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-131-909A-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-018-103A-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 7, Application US/10018103A; Publication No. US20030045465A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                          2 KHKHKHKHKGKHKHKHKHK 20
                                                                                                                                                                                                                                                             1 КНКНКНКСКНКНКНК 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM: Artificial Sequence
```

ë

ઠ 요

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 57.1%; Score 104.5; DB 15; Length 1199; Best Local Similarity 67.9%; Pred. No. 0.0009; Matches 19; Conservative 1; Mismatches 7; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match

57.1%; Score 104.5; DB 14; Length 1199;
Best Local Similarity 67.9%; Pred. No. 0.0009;
Matches 19; Conservative 1; Mismatches 7; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                     Score 110.5; DB 15; Length 1043;
Pred. No. 0.00018;
1; Mismatches 4; Indels 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-10-147-268-4
; Sequence 4, Application US/10147268
; Sequence 4, Application US/10147268
; Publication No. US20020143154A1
; GENERAL INFORMATION:
    TITLE OF INVENTION: CIF130 INHIBITS CELL CYCLE PROGRESSION
; FILE REFERENCE: PP-01513.104/200130.456D3
; FILE REFERENCE: PP-01513.104/200130.456D3
; CURRENT FILLING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; LENGTH: 1199
; TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 4, Application US/10338279
; Publication No. US20030113791A1
; GENERAL INFORMATION:
; APPLICANT: Kaufmann. Joerg
† TITLE OF INVENTION: CIF130 INHIBITS CELL CYCLE PROGRESSION
; FILE REFERENCE: PP-01513.104/200130.456D3
; CURRENT APPLICATION NUMBER: US/10/338,279
; CURRENT FILING DATE: 2003-01-07
; NUMBER OF SEQ ID NOS: 4
; SOSTWARE: FRSESEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 КН----КНКНКНКСКНКНКНКНКС-----КНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR FILING DATE: 2001/09/36
PRIOR FILING DATE: 2001-08-10
PRIOR FILING DATE: 2001-08-10
PRIOR FILING DATE: 2001-09-26
PRIOR FILING DATE: 2001-09-26
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FASESEQ for Windows Version 4.0
SEQ ID NO 258
LENGTH: 1043
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 НИНКНИКСКНИНКНИКНКНИКНК 29
                                                                                                                                                                                                                                                                                                                                                                              Query Match 60.4%;
Best Local Similarity 59.0%;
Matches 23; Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; LENGTH: 1199
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-338-279-4
                                                                                                                                                                                                                                                                                                              ; ORGANISM: Homo sapiens
US-10-097-340-258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: Homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-10-147-268-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQ ID NO 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification, TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification, TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer FILE REFERENCE: MRI-030 CURRENT APPLICATION NUMBER: US/10/097,340 CURRENT FILING DATE.
                                                                                                                                                                                                                                                                                                                                                                APPLICANT: EXELISION.
APPLICANT: EXELISION.
TITLE OF INVENTION: hPRP4s AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE FILE REFERENCE: EXO2-066C
CURRENT APPLICATION NUMBER: US/10/211,133
CURRENT APPLICATION NUMBER: US 69/310,362
PRIOR FILING DATE: 2001-08-06
PRIOR FILING DATE: 2001-08-06
PRIOR FILING DATE: 2002-02-15
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PATENTIN VERSION 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3,
                                                                                 .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ch 60.4%; Score 110.5; DB 15; Length 1007; 1 Similarity 59.0%; Pred. No. 0.00018; 23; Conservative 1; Mismatches 4; Indels 11;
                                                                                    ö
                                     Length 29;
                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 КН----КНКНКНКСКНКНКНККС-----КНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 117; DB 15;
Pred. No. 1.6e-06;
4; Mismatches 5;
                                                                                                                                                                   3 RKKRRQRRRGKHKHKHKHKHKHK 29
                                                                                                                                     1 КНКНКНКНКСКНКНКНКНКСКНКНКК 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 258, Application US/10097340 Publication No. US20030087250A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                    Sequence 7, Application US/10211133 Publication No. US20030027230A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: John MONAHAN
APPLICANT: Manjula GANNAVARAPU
APPLICANT: Sebastian HOBESCH
APPLICANT: Shubhangi KAMATYAR
APPLICANT: Steve G. KOVATS
APPLICANT: Rachel B. MEYERS
APPLICANT: Michael MORRISEY
APPLICANT: Peeer OLANDT
APPLICANT: Peeer OLANDT
APPLICANT: Peeer OLANDT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Karen LU
Rosemarie SCHMANDT
Xumei ZHAO
Karen GLATT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gordon B. MILLS
Robert C. BAST, Jr.
                                          Query Match
Best Local Similarity 66.7%;
Matches 18; Conservative '
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Peter VEIBY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: PRT;
ORGANISM: Homo sapiens
US-10-211-133-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 23; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
US-10-097-340-258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQ ID NO 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

Gaps

```
GENERAL INFORMATION:

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Hinkle, Gregory J.

APPLICANT: Gldman, Barry S.

APPLICANT: Gldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

FILE REFERENCE: 38-10(52052) B.

CURRENT APPLICATION NUMBER: US/10/369,493

CURRENT APPLICATION NUMBER: US 60/360,039

PRIOR PILING DATE: 2003-02-28

PRIOR FILING DATE: 2003-02-28

PRIOR FILING DATE: 2003-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 1406

LENGTH: 980

TYPE: PRT

ORGANISM: Saccharomyces cerevisiae

US-10-369-493-1406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 52.5%; Score 96; DB 12; Length 980; Best Local Similarity 83.3%; Pred. No. 0.0061; Matches 15; Conservative 2; Mismatches 1; Indels
2 HKHKHKHKGKHKHKHKGKKKHKHK 29
| | | | : | | | | | | | | | | | | | 1118
11148 HHHHHEHK-KKKKKKHKHKHKHKHKHKHKHKHSK 1174
                                                                                                                                      RESULT 15
US-10-369-493-1406
                                              a
  ઠે
```

Search completed: January 20, 2004, 18:28:42 Job time : 32.7826 secs

8 8

41.540	43	- √s 100 € . 3	H disk disk		E France	10 mg 1 mg							
									•				
								₽.					
		J				•							
• "					· · · .								
													•
											•	•	
		1.0											
				•									
				*		ý.							
					*								
						-					•		9
					i i								
				-00									
	•												
	• •												
		× ,											
		v											
		· ·											
				-									
•													
				·								÷	
			•										• •
									•				
			•		<u>.</u>	•							
· .	9				, *	¥							
-1-					L.	:							
				- ×	÷ .	· · · · · · · · · · · · · · · · · · ·		- A	1.3				· 2.
		· ** -		i j			7 D B						24.0
			4.					·		•		***	
			and the last	See .	Asset		Ser du	il.			interior of the		

```
GenCore version 5.1.6 (c) 1993 - 2004 Compugen Ltd.
               Copyright
```

 protein search, using sw model OM protein

January 20, 2004, 18:18:04 ; Search time 14.7101 Seconds Run on:

(without alignments) 189.590 Million cell updates/sec US-10-018-103A-4

183 1 КНКНКНКККККНКНКНККККНКНКНК 29 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5

Scoring table:

Total number of hits satisfying chosen parameters:

283308 seqs, 96168682 residues

Searched:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description	adhesin homolog HI	BEM1 protein-bindi	histidine-rich gly	ARS-binding protei	probable zinc tran	trophozoite antige	cold stress protei	hypothetical prote	 hypothetical prote 	DNA topoisomerase	histidine-rich pro	hypothetical prote	DNA topoisomerase	DNA topoisomerase	gene pipsqueak pro	pipsqueak	hypothetical prote			o,	trophozoite antige	hypothetical prote	probable membrane	kininogen, HMW pre	trophozoite antige	hypothetical prote			
	ID	D64049	S45444	KGZQHL	833791	A84696	A44863	T07078	C85838	T28937	ISHUT1	A54523	T13429	JU0144	A49546	S66148	S66149	T00063	T04219	E86291	872366	C44863	T43145	AD0262	KGHUH1	D44863	AB2396	S04491	H90992	A36664
	ОВ	~	-1	-	~	~	~	~	~	~	-	~	N	~	~	~	7	~	~	~	~	~	~	~	-	~	~	~	~	~
	Length	337	980	351	496	385	668	102	283	1257	765	140	744	167	167	532	1085	1388	658	448	829	144	356	373	644	65	353	213	279	629
ا اع مد		62.3	52.5	51.9	50.8	49.7	47.0	46.4	45.9	45.9	45.6	45.4	45.1	45.1	45.1	44.5	44.5	44.5	44.3	44.0	44.0	43.7	43.7	43.7	43.7	43.2	43.2		42.6	42.6
	Score	114	96	95	93	91	98	85	84	84	83.5	83	82.5	82.5	82.5	81.5	81.5	81.5	81	0	80.5	80	80	80	80	79	79	78	78	78
Dogill	No.	-	7	m	4	2	9	7	60	σ	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26		28	29

hypothetical prote proliferation pote nickel insertion a	urease accessory p knob protein - mal knob-associated hi	knob protein precu histidine-rich pro knob-associated Hi probable metal tra	PQ-rich protein - hypothetical prote kininogen, HMW II	kininogen, HMW I p hypothetical prote filaggrin B - mous
T50609 T42727 T51324	AD0325 A26480 A54494	A54495 A28412 B71623 D83483	S58222 T00329 KGBOH2	KGBOH1 G83138 B35026
0 0 0	000	0000	777	7 7 7
529 1560 115	231 270 473	634 634 326	400 1095 619	621 196 250
42.3 42.3	4 4 4 2 2 2 4 1 1 1 1 1	4 4 4 2 2 4 4 1 1 1 1 1 1 1	41.5 41.5 41.3	41.3 41.0 41.0
77.5	: : : : : : : : : : : : : : : : : : : :	77 77 77	76 76 75.5	75.5 75 75
30	3 8 8 8 5 4 8	8 8 4 8 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	4 4 4 0 1 2	4 4 4 6 4 6

ALIGNMENTS

4	
-	o
3	9
ű	4

C;Species: Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Accession: D64049
R;Flaischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, .; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, .; D.M.; Brandon, R.C.; Fine, L.D.; Fritchmann, J.L.; Puhrmann, J.L.; Geoghagen, N.S.M.; Science 269, 465-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Status: nucleic acid sequence not shown; translation not shown
A;Residues: 1-337 <TIGR>
A;Residues: 1-337 <TIGR>
A;Residues: 1-337 <TIGR>
A;Cross-references: GB:U32698; GB:L42023; NID:G3212178; PIDN:AAC21794.1; PID:G1573074;

Gapa ö Query Match 62.3%; Score 114; DB 2; Length 337; Best Local Similarity 60.7%; Pred. No. 4.4e-06; Matches 17; Conservative 3; Mismatches 8; Indels

ö

129 HKHEHKHDHKHDHDHDHDHKHEHKHDHE 156 2 НЮНИНКЕКНИНКНИКНИКНИКНИК 29 ð g

EEM1 protein-binding protein BOB1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YBL0717; protein YBL085w
C;Species: Saccharomyces cerevisiae
C;Date: 09-Aug-1994 #sequence revision 09-Sep-1994 #text_change 21-Jul-2000
C;Accession: 845444; 845421; \$45826; S59218
R;Bender, A.; Bender, L.; Kokojan, V.
Bubmitted to the EMBL Data Library, April 1994
A;Description: Yeast Boblp (Bemlp-binding protein) binds to the SH3 domain-containing p:
A;Reference number: 845444
A;Accession: \$45444

A;Molecule type: DNA
**Residues: 1-980 <BEND.
**Residues: 1-980 <BEND.
**A;Residues: 1-980 <BEND.
**A;Residues: 1-980 <BEND.
**A;Cross-references: EMBL:L31406; NID:g829041; PIDN:AAB08439.1; PID:g466436
**R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.
**R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.
**R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.
**A;Obermaier, Squence analysis of a 78,6 kb segment of the left end of Saccaromyces c
**A;Reference number: S45387

A;Accession: 345421 A;Molecule type: DNA A;Residues: 1-980 <DBS> A;Cross-references: EMBL:X79489; NID:g496661; PIDN:CAA56021.1; PID:g496694

Gaps

ö

```
Ag4696

Reprobable zinc transporter [imported] - Arabidopsis thaliana
Cybbable zinc transporter [imported] - Arabidopsis thaliana
Cybbable zinc transporter [imported] - Arabidopsis thaliana
Cybbate: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
Cybate: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
Cybate: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
Rilin, X.; Kaul, S.; Rounaley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
Rilin, X.; Kaul, S.; Rounaley, S.D.; Shea, J.S.; Dinayam, L.; Tallon, L.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Residues: DNA
A;Residues: DNA
A;Residues: DNA
A;Residues: 1-385 < CTO>
A;Cross-references: GB:AE002093; NID:g3980394; PIDN:AAC95197.1; GSPDB:GN00139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         trophozoite antigen R45 - malaria parasite (Plasmodium falciparum) (fragments)

C;Species: Plasmodium falciparum

C;Species: 21-Mar-1993 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C;Accession: A44863; B24863; S27832

C;Accession: A44863; B44863; S27832

Exp. Parasitol. 74, 441-451, 1992

A;Title: Plasmodium falciparum: characterization of gene R45 encoding a trophozoite ant A;Reference number: A44863; MUID:92278053; PMID:1350536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Accession: A44863
A;Molecule type: DNA
A;Residues: 1-32 x GDNI>
A;Residues: 1-32 x GDNI>
A;Cross-references: EMBL:M83793; NID:g160619; PIDN:AAA29739.1; PID:g552229
A;Experimental source: Uganda Palo Alto FUP/CB strain
                            ARS-binding protein - yeast (Kluyveromyces marxianus)
NyAlternate names: ARS-binding factor
Cjopedes: Kluyveromyces marxianus
Cjopedes: Kluyveromyces marxianus
Cjobedes: Kluyveromyces marxianus
Cjobete: OB-Jun-1994 #sequence_revision O1-Dec-1995 #text_change 29-Oct-1999
CjAccession: S33791
Rjoberye, E.H.H.; Maurer, K.; Mager, W.H.; Planta, R.J.
Biochim. Biophys. Acta al.73, 233-236, 1993
Ajritle: Structure of the ABF1-homologue from Kluyveromyces marxianus.
Ajritle: Structure of the ABF1-homologue from Kluyveromyces marxianus.
Ajritle: Structure of the ABF1-homologue from Kluyveromyces marxianus.
Ajritle: Structure of the ABF1-homologue from Kluyveromyces Ajritle: Structure of the ABF1-homologue from Kluyveromyces marxianus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Cross-references: EMBL;Z19865; NID:g312693; PIDN:CAA79673.1; PID:g312694
C;Keywords: DNA binding; nucleus; phosphoprotein; transcription regulation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 385;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 496;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
50.8%; Score 93; DB 2; Le
Best Local Similarity 34.5%; Pred. No. 0.0013;
Matches 10; Conservative 16; Mismatches 3;
protein - yeast (Kluyveromyces marxianus)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           49.7%; Score 91; DB 2;
50.0%; Pred. No. 0.0017;
iive 2; Mismatches 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       296 QHQHQHQHQHQHQHQHQSQDQHQNQHQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 КНКНКНКНКСКНКНКНКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  192 HNHSHHHHHHHHHHHKHQHQHHHK 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2 НХНКНКНКСКНКНКНКНКНКНК 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 50.0 Matches 13; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Molecule type: DNA
A;Residues: 333-668 <BON2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Accession: B44863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C, Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             · 요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                            R;Domdey, H.; Gassenhuber, H.; Obermaier, B.; Piravandi, E. submitted to the Protein Sequence Database, August 1994

A;Reference number: $45816

A;Reseasion: $45836

A;Molecule type: DNA

A;Residues: 1-980 cDNA

A;Cossa-references: BMB: $23846; NID:g536137; PIDN:CAA84906.1; PID:g536138; GSPDB:GN000C

A;Cross-references: BMB: $23846; NID:g536137; PIDN:CAA84906.1; PID:g536138; GSPDB:GN000C

R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.

R;Accession: S59218

A;Accession: S59218

A;Accession: S59218

A;Residues: 1-980 cOBW>

A;Residues: 1-980 cOBW>

A;Residues: 1-980 cOBW>

A;Rosidues: 1-980 cOBW>

A;Rosidues: 1-980 coBW>

A;Rosidues: Lhe nucleotide sequence was submitted to the EMBL Data Library, May 1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          histidine-rich glycoprotein precursor - Plasmodium lophurae
C;Species: Plasmodium lophurae
C;Species: Plasmodium lophurae
C;Date: 30-Sep-1997 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C;Date: 30-Sep-1897 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
R;Ravetch, J.V.; Feder, R.; Pavlovec, A.; Blobel, G.
R;Ravetch, J.V.; Feder, R.; Pavlovec, A.; Blobel, G.
Nature 312, 616-620, 1984
A;Title: Primary structure and genomic organization of the histidine-rich protein of the A;Reference number: A22692; MUID:85061618; PMID:6095114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Molecule type: DNA
A;Residues: 1-351 <RAV>
A;Cross-references: GB:X01469; NID:g9997; PIDN:CAA25698.1; PID:g9999
C;Comment: There are two copies of 16-residue repeats, two copies of 17-residue repeats,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Map position: 2L
C;Superfamily: BEM1 protein-binding protein BOB1; pleckstrin repeat homology; SAM homold
F;20-72/Domain: SH3 homology <SH3>
F;225-291/Domain: SAM homology <SAM>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gapa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Introns: 23/3
A;Introns: 23/3
A;Introns: 23/3
A;Introns: 23/3
C;Superfamily: plasmodium histidine-rich protein
C;Reywords: glycoprotein; tandem repeat
F;1-23/Domain: signal sequence #status predicted <SIG>F;2-47/Domain: propeptide #status predicted <PRO>F;48-351/Product: histidine-rich glycoprotein #status predicted <PRO>F;59-74,75-90/Region: 16-residue repeats
F;31-107,108-113/Region: 11-residue repeats
F;124-138,139-153/Region: 15-residue repeats
F;173-301,312-331/Region: 10-residue repeats
F;40/Binding site: carbohydrate (Asn) (covalent) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 51.9%; Score 95; DB 1; Length 351; Best Local Similarity 51.9%; Pred. No. 0.00058; Matches 14; Conservative 0; Mismatches 13; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       52.5%; Score 96; DB 1; Length 980; 83.3%; Pred. No. 0.0011; 1ive 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C;Genetics:
A;Gene: SGD:BOI1; BOB1; MIPS:YBL085w
A;Cross-references: MIPS:YBL085w; SGD:S000181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   215 нинининенининининенинин 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EGKHKHKHKHKSKHKHKN 702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9 КСКНКНКНКНКСКНКНКН 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 83.33
Matches 15; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ដ
```

ö

Gaps

ö

RESULT 4 S33791

ო

```
2 НКНКНКНКСКНКНКНКНККККНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Match 45.9%;
Local Similarity 55.2%;
He 16; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A, Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Accession: S13821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 10
                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C;Species: Bacherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: C85838
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, Iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7.
A;Reference number: A85480; MUD:21074935; PMID:11206551
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hypothetical protein yohm [imported] - Escherichia coli (strain O157:H7, substrain EDL93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A,Cross-references: GB:AE005174; NID:g12516319; PIDN:AAG57167.1; GSPDB:GN00145; UWGP:Z32
A,Experimental source: strain O157:H7, substrain EDL933
A;Cross-references: GB:M83792; NID:g160617; PIDN:AAA29738.1; PID:g160618
A;Experimental source: Uganda Palo Alto FUP/CB strain
A;Note: sequence extracted from NCBI backbone (NCBIP:105026) and corrected to correspond
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                         C;Keywords: glycoprotein; surface antigen; tandem repeat
F;225-329/Region: 6-residue repeats (H.K.S-D-S/H-N)
F;373-379/Region: 6-residue repeats (H.K.S-D-S/H-N)
F;222/Binding site: carbohydrate (Asn) (covalent) #status predicted
                                                                                                                                                                                                                                                                 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 45.9%; Score 84; DB 2; Length 283; Best Local Similarity 42.9%; Pred. No. 0.008; Matches 12; Conservative 3; Mismatches 13; Indels
                                                                                                                                                                                                                      Length 668;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 102,
                                                                                                                                                                                              47.0%; Score 86; DB 2; Length 668
50.0%; Pred. No. 0.0099;
mismatches 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 85; DB 2; Length 102
Pred. No. 0.0026;
6; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                 349 HKSDHNHKSDHNHKSDNNHKSDHNHKSDHK 378
                                                                                                                                                                                                                                                                                                         2 НИНКНИНКСКНИНК -- НКНКСКНИНКНИК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 КНКНКНКНКСКНК--НКНКНКСКНКНКНК 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C;Genetics:
A;Gene: src1
C;Superfamily: cold stress protein COR19
C;Keywords: cold shock
                                                                                                                                                                                                                                                                                                                                                                                                                                                             cold stress protein SRC1 - soybean
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 51.7%;
Matches 15; Conservative 6
                                                                                                                                                                                                                    Query Match
Best Local Similarity 50.0°
Matches 15; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: DNA
A; Residues: 1-283 <STO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Gene: yohM
                                                                                         A;Gene: R45
                                                                   Genetics:
                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
```

```
DNA topoisomerase (EC 5.99.1.2) - human
N;Alternate names: nicking-closing enzyme; omega-protein; relaxing enzyme; swivelase; t;
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1992 #sequence revision 30-Sep-1992 #text_change 18-Jun-1999
C;Accession: A30887; A40008; Si1821; S02397; S40643; A34422
R;D'Arpa, P.; Machlin, P.S.; Ratrie III, H.; Rothfield, N.F.; Cleveland, D.W.; Barnshaw A;Title: cDNA cloning of human DNA topoisomerase I: catalytic activity of a 67.7-kDa ca-A;Reference number: A30887; MUID:88190108; PMID:2833744
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Molecule type: mRNA
A; Residues: 1-765 cDAR.
A; Residues: 1-765 cDAR.
A; Residues: 1-765 cDAR.
A; Cross-references: GB:J03250; NID:g339805; PIDN:AAA61207.1; PID:g339806
A; Title: Structure of the human type I DNA topoisomerase gene.
A; Title: Structure of the human type I DNA topoisomerase gene.
A; Reference number: A40008; MUID:91236733; PMID:1851751
A; Residues: 1-144, A', 146-553, 'E', 555-765 <KUN>
A; Molecule type: DNA
A; Molecule type: DNA
A; Residues: 1-144, 'A', 146-553, 'E', 555-765 <KUN>
A; Cross-references: GB:M60688; GB:M60689; GB:M60690; GB:M60691; GB:M60692; GB:M60705; G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: DNA
A; Residues: 1-20 (AUJ2)
A; Residues: 1-20 (AUJ2)
A; Cross-references: EMBL:X52601
A; Cross-references: EMBL:X52601
Bur. J. Biochem. 177, 521-529, 1988
A; Trile: Monoclonal antibodies neutralizing mammalian DNA topoisomerase I activity.
A; Reference number: S02397; MUID:89064806; PMID:2461859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Cross-references: EMBL:U64598; PIDN:AAC47974.1; GSPDB:GN00028; CESP:C52B9.8
A;Experimental source: strain Bristol N2; clone C52B9
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                     C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 84; DB 2; Length 1257;
Pred. No. 0.028;
1; Mismatches 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Gene: CESP:C52B9.8
A;Map position: X
A;Introns: 15/2; 321/3; 450/3; 596/2; 776/2; 823/2; 871/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C;Accession: T28937
R;Nelson, J.
Bubmitted to the EMBL Data Library, July 1996
A;Description: The sequence of C. elegans cosmid C52B9.
A;Reference number: Z20545
A;Accession: T28937
                                                                                                                                                                                                                                                                                                                      hypothetical protein C52B9.8 - Caenorhabditis elegans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1257 <NEL>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1152 KEPHKEKEKDKEKEKEDKGKEKEKHKDK 1180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 КНКНКНКНКСКНКНКНКНКНКККНКНК 29
124 HHEDHEHHODHDHDHDHEHKEHE 151
```

```
DNA topoisomerase (EC 5.99.1.2) - mouse
NyAlternate names: type I DNA topoisomerase
NyAlternate names: type I DNA topoisomerase
C;Species: Mus musculus (house mouse)
C;Species: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Jun-2000
C;Accession: JU0144
R;Koiwai, O; Yasui, Y.; Sakai, Y.; Watanabe, T.; Ishii, K.; Yanagihara, S.; Andoh, T.
A;Title: Cloning of the mouse cDNA encoding DNA topoisomerase I and chromosomal location
A;Reference number: JU0144; MUID:93216125; PMID:8096488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A49546
DNA topoisomerase (EC 5.99.1.2) - Chinese hamster
c;Species: Cricetulus griseus (Chinese hamster)
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 16-Jul-1999
C;Accession: A49546; 832697; $32698
C;Accession: A49546; Bartrand, R.; Kohlhagen, G.; Tabuchi, A.; Jenkins, J.; Pommier, Y.
J; Ranizawa, A.; Bertrand,
J; Biol. Chem. 268, 25463-25468, 1993
A;Title: Cloning of Chinese hamster DNA topoisomerase I cDNA and identification of a si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A,Molecule type: mRNA
A,Residues: 1-767 -KOL2.
A,Gross-references: GB:Dl0061; NID:g220617; PIDN:BAA00950.1; PID:g220618
C,Cross-references: GB:Dl0061; NID:g220617; PIDN:BAA00950.1; PID:g220618
C,Comment: This enzyme catalyzes the transient breakage and rejoining of single strande
                                                                                                                                                                                                                                                                                                                                                                              A,Cross-references: EMBL:AL009195; NID:e1355203; PID:e1202214; PIDN:CAA15704.1
C,Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----KGKH------KHKHKHKGKH------KHKH 26
                       C;Species: Drosophila melanogaster

C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000

C;Accession: T13429

R;Auxphy, L; Harris, D.; Barrell, B.

R;Muxphy, L; Harris, D.; Barrell, B.

R;Muxphy, L; Harris, Data Library, April 1999

A;Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A;Reference number: Z17668

A;Accession: T13429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22 HKHKDKHKDREHRHKEHKKDKDREKSKHSBHKDSEKKHKEKEKTKHKDGSSEKHKD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 767;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 744;
hypothetical protein 30B8.6 - fruit fly (Drosophila melanogaster)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  h 45.1%; Score 82.5; DB 2; Similarity 34.9%; Pred. No. 0.027; 22; Conservative 0; Mismatches 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 2;
                                                                                                                                                                                                                                                                                       A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-744 <MUR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A,Map position: 2-54.5
C;Superfamily: eukaryotic type I DNA topoisomerase
C;Keywords: DNA binding; isomerase
F;725/Active site: Tyr #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0.027;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 82.5; DE Fred. No. 0.027 8; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            130 НЅНННОНОНОНОНОНОНОНОН---НОНКН 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 нкникникскнинкникнин 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Cross-references: FlyBase: FBgn0000377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      45.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 HKHKHKH------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 44.0°
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Map position: X
A; Introns: 49/3; 178/3; 625/1
A; Note: EG:30B8.6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27 KHK 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Accession: JU0144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  XHX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Gene: TopI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C, Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Rirving, D.O.; Cross, G.A.M.; Feder, R.; Wallach, M.
Mol. Biochem. Parasitol. 18, 223-234, 1986
A;Title: Structure and organization of the histidine-rich protein gene of Plasmodium lop
A;Reference number: A54523; MUID:86174893; PMID:3007981
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: mRNA
A; Residues: 657-765 <MAU>
A; Crose-references: GB:M27913; NID:g339807; PIDN:AAA61208.1; PID:g339808
A; Crose-references: GB:M27913; NID:g339807; PIDN:AAA61208.1; PID:g339808
C; Comment: Type I DNA topoisomerase catalyzes the ATP-independent transient breakage of in another, followed by rejoining. This reaction will lead to the conversion of one top
                                                                                                                                                                                                                                                                                                                                                                                                     ă
                                 A,Residues: 344-765 <ODD>
A,Cross-references: GB:M60657
A;Cross-references: GB:M60657
S,Tamura, H.O.; Kohchi, Yamada, R.; Ikeda, T.; Koiwai, O.; Patterson, E.; Keene, J.L
Nucleic Acids Res. 19, 69-75, 1991
A;Title: Molecular cloning of a cDNA of a camptothecin-resistant human DNA topolsomerase
A;Reference number: $40643; MUID:91187651; PMID:1849260
                                                                                                                                                                                                                                                                                          A;Molecule type: mRNA
A;Residues: 523-543;573-582,'D',584-593 <TAM>
R;Maul, G.G.; Jünenez, S.A.; Riggs, E.; Ziemnicka-Kotula, D.
Proc. Natl. Acad. Sci. U.S.A. 86, 842-8496, 1989
A;Title: Determination of an epitope of the diffuse systemic sclerosis marker antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HXHKHKHKH------KGKH------KGKKH------KHKHKKKKK 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22 HKHKDKHKDREHRHKEHKKEKDREKSKHSNSEHKDSEKKHKEKEKTKHKDGSSEKHKDKH 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             histidine-rich protein - Plasmodium lophurae (fragment)
C;Species: Plasmodium lophurae
C;Date: 15-Oct_1994 #sequence_revision 15-Oct-1994 #text_change 20-Aug-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 33; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A,Accession: A54523
A,Status: preliminary
A,Molecule type: mRNA
A,Residues: 1-140 <IRV>
A,Cross-references: GB:MIS317; NID:g160331; PIDN:AAA29616.1; PID:g552196
C,Superfamily: plasmodium histidine-rich protein
C,Keywords: tendem repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Cross-references: GDB:120444; OMIM:126420
A;Map position: 20q12-20q13.1
A;Introns: 11/3; 20/1
A;Note: the list of introns is incomplete
C;Superfamily: eukaryotic type I DNA topoisomerase
C;Keywords: DNA binding; DNA replication; heterotetramer; isomerase
F;723/Active site: Tyr #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              45.6%; Score 83.5; DB 1; Length 765; 36.1%; Pred. No. 0.021; tive 0; Mismatches 6; Indels 3:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          / Match 45.4%; Score 83; DB 2; Length 140; Local Similarity 44.4%; Pred. No. 0.0057; nes 12; Conservative 0; Mismatches 15; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Reference number: A34422; MUID:90046823; PMID:2479024
A;Accession: A34422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              нинининарнинининин 35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 НКНКНКНКСКНКНКНКНКНКН 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 36.1<sup>§</sup>
Matches 22; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ity in systemic sclerosis.
                                                                                                                                                                                                                                                                Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Accession: A54523
                                                                                                                                                                                                                                  A; Accession: S40643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29 K 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              82 K 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Gene: GDB: TOP1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
T13429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
```

셤 ò g

ò

Gaps

3,

35;

81

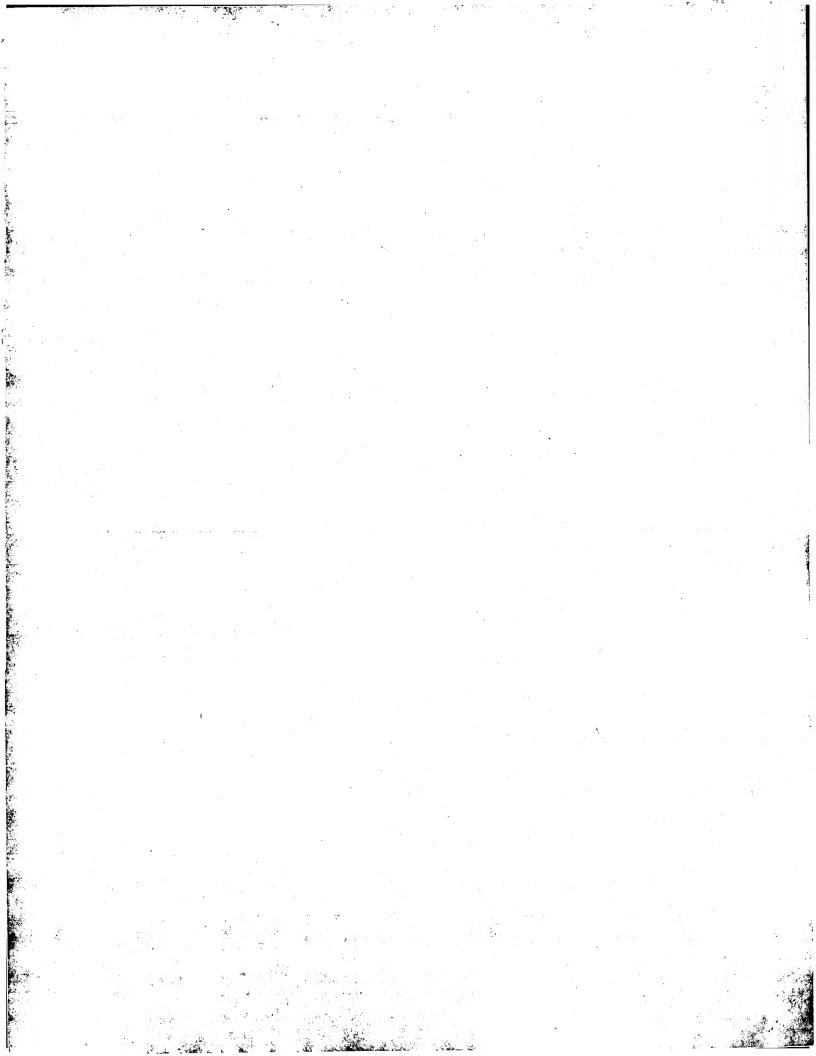
ઠ 셤

A; Reference number: A49546; MUID: 94064611; PMID: 8244980

```
gene pipsqueak protein A short form - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Species: Drosophila melanogaster
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jul-2000
C;Accession: S6648
R;Weber, U.; Siegel, V.; Mlodzik, M.
BEMBO J. 14, 6247-6257, 1995
A;Title: pipsqueak encodes a novel nuclear protein required downstream of seven-up for A;Reference number: S66148; MUID:96134923; PMID:8557044
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ñ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 HKHKHKH-----KGKH------KGKH-----KHKHKHKGKH-----KHKH 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22 HKHKDKHKDREHRHKEHKKDKEKDREKSKHSNSEHKDSEKKHKEKEKTKHKDGSSEKHKD 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A,Status: preliminary
A,Molccule type: DDA
A,Residues: 1-535 cMEB>
A,Cross-references: EMBL:X90986; NID:g1149498; PIDN:CAA62473.1; PID:g1149499
                                                                                                                                                                                                                                                                                              A; Molecule type: mRNA
A; Residues: 1-504, 8', 506-767 < TA2>
A; Crose-references: EMBL: Z21624; NID: 9297076; PIDN: CAA79747.1; PID: 9297077
A; Experimental source: DC3F/C-10 cells
A; Note: this form is camptothecin-resistant
C; Superfamily: enkaryotic type I DNA topoisomerase
C; Keywords: isomerase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 34.9%; Pred. No. 0.027;
Matches 22; Conservative 0; Mismatches 6; Indels 35; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
A;Molecule type: mRNA
A;Residues: 1-767 <TAN>
A;Residues: 1-767 <TAN>
A;Crosg-references: GB:221625; NID:g297078; PIDN:CAA79748.1; PID:g297079
A;Experimental source: DC3F cells
A;Note: this form is not camptothecin-resistant
B;Tanizawa, A; Tabuchi, A.; Bertrand, R.; Pommier, Y.
submitted to the EMBL Data Library, Pebruary 1993
A;Reference number: S32697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
44.5%; Score 81.5; DB 2; Length 535;
Best Local Similarity 39.4%; Pred. No. 0.026;
Matches 13; Conservative 4; Mismatches 11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 НКНКИКН-----КСКНКНКНКНКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C;Genetics:
A;Gene: pipsqueak
C;Superfamily: POZ domain homology
F;21-123/Domain: POZ domain homology <POZ>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           27 KHK 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        82 KHK 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           名
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
```

셤

Search completed: January 20, 2004, 18:24:42 Job time : 15.7101 secs



Н

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

OM protein - protein search, using sw model

January 20, 2004, 18:06:29; Search time 9.66667 Seconds (without alignments) 141.080 Million cell updates/sec Run on:

US-10-018-103A-4 183 1 KHKHKHKHKHKHKHKKKKKKKK 29 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues Searched:

127863 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_41:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Þt	P44526 haemophilus	~		_	_			P72204 pasteurella		Q9xt26 equus cabal						Q9pub8 brachydanio							9		Q9v3a4 drosophila		4	P11088 mus musculu	P49907 bos taurus	Q9xuc4 caenorhabdi	w	60583	P50439 caenorhabdi
Saranas	ID	ZNUA HAEIN	PR4B_HUMAN	PR4B_MOUSE	BOB1_YEAST	HRPX_PLALO	BAF1 KLUMA	HYPB_BRAJA	TONB PASHA	CCT1 HUMAN	CCT1_HORSE	TOP1 HUMAN	TOP1_CRIGR	TOP1 MOUSE	T2D2_DROME	TOP1_XENLA	KE4_BRARE	KNG_HUMAN	удня вснро	SKGR_XENLA	HMN1_DROME	UREE YERPE	KNOB PLAFA	KNOB_PLAFG	UREE YEREN	CSUP_DROME	KNH2_BOVIN	KNH1_BOVIN	FILA MOUSE	SELP BOVIN	YGJK_CAEEL	SELP_RAT	CCT2 HUMAN	YV59_CAEEL
	08	-	-	-	-	-	-	-	-		-	-	-	ч	н	Н	Н	-	-	-	-	Н	-	-	-	-	-1	-	٦	-		Н	Н	н
	Length	:	1001	1007	980	351	496	302	246	726	727	765	767	767	_		352	644	732	213	629	231	473	634	230	449	619	621	336	402	404	385	730	147
#	Query	62.	60.4	60.4	52.5	51.9	50.8	46.4	46.2	45.9	45.9	45.6	45.1	45.1	44.5	44.0	43.7	43.7	43.7	42.6	42.6	42.1	42.1	42.1	41.3	41.3	41.3	41.3	41.0	41.0	41.0	40.7	40.7	40.4
	Score		110.5	110.5	96	95	93	85	84.5	84	m		82.5			80.5	80	80	80	78	78	77	77	7	75.5	75.5	75.5		75	75	75		•	74
	Result No.	-	7	e	4	S	9	7	89	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	30	31	32	33

P52317 yersinia ps P08016 schistosoma Q64733 mus musculu P14586 plassmodium P1458 plassmodium P31705 candida alb P53705 candida alb P6719 plassmodium P20105 drosophila P11536 drosophila P11536 drosophila P11536 drosophila	OS4620 muscura
UREE YERPS EGGS SCHWA FYB2 MOUSE HRP3 PLAFS HRP9 ALCEU INT1 CANAL YKR1 CAREL KNOB PLAFN E74A DROME B74B DROME A710 HUMAN	AFTO_MOOSE
алалалалал	4
231 149 428 428 361 1664 886 657 829 883	1000
4 4 4 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	0.00
74 73.5 73.5 72.5 72.5 71.5 71.7	7
W W W W W W A A A A A A A A A A A A A A	ņ

ALIGNMENTS

```
EMBL; AY029347; AAK38155.1; -.
EMBL; AF283465; AAM19101.1; -.
EMBL; AB011108; BAA25462.1; ALT_INIT.
EMBL; A48736; AAB03268.1; -.
HSSP; P24941; 1AQ1.
Genew; HGNC:17346; PRPF4B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
MEDLINE=22072586; Pubmed=12077342;
                                                                                                                                                                                                                                                                                                                                                                                                      60.4%;
                                                                                                                                                                                                                                                                                                                                                                                   116972
                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 59.0 tes 23; Conservative
                                                                                                                                                                                                                                                                                                                       78
496
1006
701
                                                                                                                                                                                                                                                                                                                                                                         815
                                                                                                                                                                                                                                                                                                                                                                       815 81
1007 AA;
                                                                                                                                                                                                                                                                                                            Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                   693
                                                                                                                                                                                                                                                                                                                                                                          ACT SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                     NP BIND
BINDING
                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PR4B MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   요
                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      013523; OGTDP2; O960T7; O9UEB6; 15-JUL-1998 (Rel. 36, Created) 28-FEB-2003 (Rel. 41, Last sequence update) 28-FEB-2003 (Rel. 41, Last annotation update) serine/threonine-protein kinase PRP4 homolog (EC 2.7.1.37) (PRP4 pre-
                   HIGH-AFFINITY ZINC UPTAKE SYSTEM PROTEIN
                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";

DNA Res. 5:31-39(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Functional analysis of the fission yeast Prp4 protein kinase involved in pre-mRNA splicing and isolation of a putative mammalian
                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.

BEDILME-22072565, PubMed=12077342;

Dellaire G., Makarov E.M., Cowger J.J.M., Longman D.,

Sutherland H.G.E., Luhrmann R., Torchia J., Bickmore W.A.;

Sutherland H.G.E., Luhrmann R., Torchia J., Bickmore W.A.;

"Mammalian RPP4 kinase copurifies and interacts with components of both the US shrNP and the N-CoR deacetylase complexes.";

MOI. Cell. Biol. 22:5141-5156 (2002).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 512-1007 FROM N.A.
MEDLINE-97250302; PubMéd-9102632;
Gross T., Lutzelberger M., Wiegmann H., Klingenhoff A., Shenoy S.,
Kaeufer N.F.;
                                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TISSUE-Brain;
MEDLINE-98290545; PubMed-9628581;
Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
Nagase T., Ishikawa C.-I., Miyajima N., Tanaka A., Kotani H.,
Nomura N., Ohara O.;
                                                                                                      Score 114; DB 1; Length 337; Pred. No. 8e-07;
                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                      SECUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE=21402934; PubMed=11418604;
Kojima T., Zama T., Wada K., Onogi H., Hagiwara M.;
"Cloning of human PRP4 reveals interaction with CIK1.";
J. Biol. Chem. 276:32247-32256(2001).
                                                   ZINC (BY SIMILARITY).
ZINC (BY SIMILARITY).
ZINC (BY SIMILARITY).
3DBB45AB8F06FFCB CRC64;
                                                                                                                             8
                                                                                                                                                                                                                        PRT; 1007 AA.
                                                                                                                            3; Mismatches
                                                                                                                                                            129 HKHEHKADHKADHCHDHDHDHKHEHKADHE 156
                                                                                                                                                2 НКНКНКНКНКНКНКНКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nucleic Acids Res. 25:1028-1035(1997).
                                            HIS-RICH
                                                                                     37659 MW;
                                                                                                          62.3%;
                                                                                                        Query Match
Best Local Similarity 60.7
                                                                                                                                                                                                                          STANDARD;
                                                                             235
                                                                                                                                                                                                                                                                                                           Homo sapiens (Human)
                                             115
57
171
235
337 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
      Complete proteome.
SIGNAL
                      24
                                                                                                                                                                                                                         PR4B HUMAN
Q13523; Q8TE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             homologue.
                                                                                        SEQUENCE
                                              DOMAIN
                          CHAIN
                                                                                                                                                                                                                PR4B_HUMAN
                                                                                                                                                                                                       RESULT 2
                                                                                                                                                                                                                            KW
FT
FT
FT
SQ
SQ
                                                                                                                                                     ઠ
                                                                                                                                                                        g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (PRP4 pre-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae; Mus.
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Dellaire G., Makarov E.M., Cowger J.J.M., Longman D., Sutherland H.G.E., Luhrmann R., Torchia J., Bickmore W.A.; Mammalian PRP4 Kinase copurifies and interacts with components both the US snRNP and the N-CoR deacetylase complexes."; Mol. Cell. Biol. 22:5141-5156 (2002)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Probom; PD000001; Proc Kinase; 1.
SMART; SM00220; S_TKC; 1.
SMART; SM00219; TyrKC; 1.
PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
PROSITE; PS00109; PROTEIN KINASE_ST; 1.
PROSITE; PS00101; PROTEIN KINASE_DOM; 1.
MRNA processing; mRNA splicing; Transferase;
serine/threonine-protein kinase; ATP-binding; Phosphorylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 110.5; DB 1; Length 1007; Pred. No. 5.6e-06;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT5 1007 AA.

PRTB MOUSE STANDARD;

Q61136; O88378; Q8R4Y5; Q9CTL9; Q9CTT0;

15-JUL-1998 (Rel. 36, Created)

28-FEB-2003 (Rel. 41, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)

28-FEB-2003 (Rel. 44, Last annotation update)

Serine/threonine-protein kinase PRF4 homolog (EC 2.7.1.37)

RRNA processing factor 4 homolog) (Pre-mRNA protein kinase)

PRPF4B OR PRPK OR PRR4H OR PRP4M OR PRP4K OR CBP143.

Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GK; Q13523; -...
MIM; 602338; -...
GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
GD:0000719; Prot_kinase.
InterPro; IPR0002290; Ser_thr_pkinase.
InterPro; IPR001245; Tyr_pkinase.
PF00069; pkinase; 1...
PFam; PF00069; pkinase; 1...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 1-279 FROM N.A.
STRAIN=C57BL/6J; TISSUE=Hippocampus, and Urinary bladder;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           45ECF73ABB56D17C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19 KHSRHKKKKHKKKHSKHSKHSEEDKDKOKHKHKHKH 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HIS-RICH.
ARG/LYS-RICH.
PROTEIN KINASE.
ATP (BY SIMILARITY).
ATP (BY SIMILARITY).
BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 KH----KHKHKHKGKHKHKHKHKG-----KHKHKHKHK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Mismatches
```

```
Matches
  55555555
                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DATA TARKARA T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce) or send an email to license@isb-sib.ch).
            Makai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Komo H., Adachi J., Fukuda S., Arakawa T., Hara A., Fukunishi Y., Komo H., Adachi J., Fukuda S., Anadawa T., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Golobori T., Bono H., Kasukawa T., Saito R., Adota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Richi B., Lewis S., Matsuo Y., Nikaido I., Ring B., Kochiwa H., Ruehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Whinhaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 185-1007 FROM N.A.
MEDILTE-9836054; PubMed=9701556;
Tate P., Lee M., Tweedie S., Skarnes W., Bickmore W.;
"Capturing novel mouse genes encoding chromosomal and other nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Punctional analysis of the fission yeast Prp4 protein kinase involved in pre-mRNA splicing and isolation of a putative mammalian
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=97250102; PubMed=9102612;
Gross T., Lutzelberger M., Wiegmann H., Klingenhoff A., Shenoy S.,
Kaeufer N.F.;
                                                                                                                                                                                                                                                                                                                                                                                                                                       Hayashizaki Y.;
"Functional annotation of a full-length mouse cDNA collection.";
Nature 409:685-690(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
PROSITE; PS00108; PROTEIN KINASE ST; 1.
PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
RINA processing; MRNA splicing; Transferase;
Serine/threonine-protein kinase; ATP-binding; Phosphorylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MGD; MGI:109584; Prpf4b.
GO; GO:0005694; C:chromosome; IDA.
GO; GO:0005694; C:chromosome; IDA.
InterPro; IPR00719; Prot kinase.
InterPro; IPR001299; Ser thr pkinase.
InterPro; IPR001245; Tyr_pkinase.
Ffam; Pr00069; pkinase; I.
ProDm; PD000001; Prot kinase; I.
SMART; SM0220; S TKC; I.
  MEDLINE=21085660; PubMed=11217851;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Cell Sci. 111:2575-2585(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; AF283466; AAM19102.1; --
EMBL; AK021274; BAB32137.1; --
EMBL; AK021274; BAB3258.1; --
EMBL; AF033663; AAG32042.1; --
EMBL; U48727; AAB03269.1; --
HSSP; P24941; IAQ1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE OF 512-1007 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nuclear protein.
```

```
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SGD; S0000181; BOIL.

GQ; GO:0005935; C:bud neck; IDA.

GQ; GO:0005543; P:phospholipid binding activity; IDA.

GQ; GO:00005843; P:setablishment of cell polarity (sensu Sacch. . .; IGI.

GQ; GO:0007266; P:Rho protein signal transduction; IGI.

InterPro; IPR001849; PH.

InterPro; IPR001660; SAM.

InterPro; IPR001452; SAM.
                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDLINE-96234675; PubMed-8666672;
Bender L., Lo H.S., Lee H., Kokojan V., Peterson V., Bender A.;
"Associations among PH and SH3 domain-containing proteins and Rho-type GTPases in Yeast.";
J. Cell Biol. 133:879-894(1996).
                                                                                                                                                                                                                                                                                    Score 110.5; DB 1; Length 1007;
Pred. No. 5.6e-06;
1; Mismatches 4; Indels 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Yeast 11:1103-1112(1995).
--- FUNCTION: BINDS TO THE BEM1 PROTEIN.
--- SIMILARITY: Contains 1 SH3 domain.
--- SIMILARITY: Contains 1 PH domain.
--- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
                  ARG/LYS-RICH.
PROTEIN KINASE.
ATP (BY SIMILARITY).
BY SIMILARITY).
SKS -> IFG (IN REF. 3).
K -> I (IN REF. 3).
F -> L (IN REF. 4).
I MW; 18E6C3C43BE7AB4C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Obermaier B., Gassenhuber J., Piravandi E., Domdey H.; "Sequence analysis of a 78.6 kb segment of the left end of Saccharomyces cerevisiae chromosome II.";
                                                                                                                                                                                                                                                                                                                                                                                                  1 KH----KHKHKHKKKKKKKKKKKKG-----KHKHKKKKK 29
                                                                                                                                                                                                                                                                                                                                                                                                                              40 KHSRHKKKKHKKKHSKHSKHSEEDRDKHKHKHK 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-OCT-1994 (Rel. 30, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
BOB1 protein (BEM1-binding protein)
BOB1 Drotein (REM1-binding protein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     980 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=96076635; PubMed=7502586;
41 79 H.
40 497 A.
687 1006 P.
693 701 A.
717 717 A.
815 815 B.
185 187 E.
223 223 ?
633 633 I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (Rel. 30, Created)
                                                                                                                                                                                                                                                                                         ch 60.4%;
1 Similarity 59.0%;
23; Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; X79489; CAAS6021.1; -. EMBL; Z35846; CAA84906.1; -. PIR; S45444; S45444.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; L31406; AAB08439.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                    Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN=S288c;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BOB1 YEAST
P38041;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-OCT-1994
        DOMAIN
DOMAIN
DOMAIN
NP BIND
BINDING
ACT SITE
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BOB1_YEAST
```

Length 351;

```
296 QHQHQHQHQHQHQHQHQSQDQHQNQHQ 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 КНКНКНКНКСКНКНКНКНКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HYPB_BRAJA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 7
                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                             ò
                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                            MEDILINE-85061618; PubMed-6095114; Blobel G.; MEDILINE-85061618; Perlowec A., Blobel G.; MEDILINE-85061618; Pavlowec A., Blobel G.; Merchary Bructure and genomic organization of the histidine-rich protein of the malaria parasite Plasmodium lophurae."; Nature 312:616-620(1984).

INSCELLANEOUS: IN THE INTRARRYTHROCYTIC STAGES OF DEVELOPMENT OF PLOPHURAE IN DUCKS, THERE IS A SYNTHESIS OF A MAJOR PROTEIN ACCUMULATES TO COMPRISE AT LEAST 50% OF THE CELLULAR MASS: THE HISTIDINE RICH PROTEIN.
                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (PROBABLE)
                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                            13-AUG-1987 (Rel. 05, Created)
13-AUG-1987 (Rel. 05, Last sequence update)
15-AUG-1989 (Rel. 38, Last annotation update)
Histidine-rich glycoprotein precursor.
Plasmodium lophurae.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                               DB 1; Length 980;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HISTIDINE-RICH GLYCOPROTEIN
                                                                                                                                                52.5%; Score 96; DB 1; Length 980
83.3%; Pred. No. 0.00025;
vative 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           B X 10 AA TANDEM REPEATS.
D19A48D47D890453 CRC64;
                                                                                                                            09F1DD1F9EF30F36 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 X 15 AA TANDEM RÈPEATS.
15-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LINKED (GLCNAC. . .) (IX 16 AA TANDEM REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 X 17 AA TANDEM REPEATS
                                                                                                                                                                                                                                                           351 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIR; A22692; KGZQHL.
Malaria; Repeat; Glycoprotein; Signal
Pfam; PF00169; PH; 1.
Pfam; PF00536; SAM; 1.
Pfam; PF00018; SH3; 1.
SWART; SW00233; PH; 1.
SWART; SW00454; SAM; 1.
PR0SITE; PS50003; PH DOWAIN; 1.
PR0SITE; PS50003; PH DOWAIN; 1.
PR0SITE; PS50002; SH3; 1.
                                                                                                SH3.
SAM.
PH.
                                                                                                                                                                2;
                                                                                                                                                                                                   EGKHIKHKHKKKKHKHKN 702
                                                                                                                                                                                        9 КСКНКНКНКНКНКН 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        44032 MW;
                                                                                                                             980 AA; 109295 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; X01469; CAA25698.1; -.
                                                                                                                                                                      15; Conservative
                                                                                                                                                                                                                                                            STANDARD;
                                                                                                 292
895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       108
124
124
139
173
351 AA;
                                                                                                                                                           Local Similarity
                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=5853;
                                                                                                  13
· 228
776
                                                                                                                                                                                                                                                            HRPX PLALO
P04929;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
DOMAIN
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
DOMAIN
SEQUENCE
                                                                                                                                                                                                             685
                                                                                                                                SEQUENCE
                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
REPEAT
REPEAT
DOMAIN
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROPEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIGNAL
                                                                                                    DOMAIN
                                                                                                                                                               Best Loc
Matches
                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                           ò
```

```
This SWISS-RROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=CBS 6556;
MEDLINE=32377959; PubMed=7916634;
MEDLINE=32377959; PubMed=7916634;
Oberye E.H.H., Maurer K., Mager W.H., Planta R.J.;
Oberye E.H.H., Maurer K., Mager W.H., Planta R.J.;
Structure of the ABRI-homologue from Kluyveromyces marxianus.";
Biochim. Biophys. Acta 1173:233-236(1993).
-1- FUNCTION: TRANS-ACTING FACTOR IN THE REGULATION OF TRANSCRIPTION
AND IN DAR REPLICATION. INVOLVED IN THE TRANSCRIPTION ACTIVATION
OF A SUBSET OF RIBOSOMAL PROTEINS GENES. BINDS THE ARS-ELEMENTS
FOUND IN MANY PROMOTERS. BINDS TO THE SEQUENCE S'-TCN(7)ACG-3'.
                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR; $33791; $33791.
InterPro; IPR006774; BAF1_ABF1.
InterPro; IPR00684; BAF1_ABF1.
Transcription regulation; Activator; DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger; Metal-binding; Zinc; DNA replication;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HIS-RICH.
GLN/HIS-RICH.
POLY-ALA.
PHOSPHORYLATION (BY PKC) (POTENTIAL)
RFIEDS60D9C2E58F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kluyveromyces marxianus (Yeast) (Kluyveromyces fragilis).
Kluyveromyces, Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Bukaryota; Fungi; Asccharomycetaceae; Kluyveromyces.
Saccharomycetales; Saccharomycetaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                    (Rel. 28, Created)
(Rel. 28, Last sequence update)
(Rel. 29, Last annotation update)
n factor BAPI (ARS binding factor 1) (Protein ABF1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -!- SIMILARITY: STRONG TO YEAST BAF1, AND LOCAL TO YEAST RAP1.
                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBCELLULAR LOCATION: Nuclear.
PTM: EXTENSIVELY PHOSPHORYLATED ON SER AND THR RESIDUES
(BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 93; DB 1; Length 496;
Pred. No. 0.00029;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3; Indels
Score 95; DB 1; I
Pred. No. 0.00012;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      16; Mismatches
                                                                                                                                                                                                                                                                                      496 AA
                                                      0; Mismatches
                                                                                                                                                    215 нининининининининининини 241
                                                                                                         2 НИНИНИНКЕКНИНИНИНКЕКНИНИНИН
                                                                                                                                                                                                                                                                                                                                                                                                                     Transcription factor BAF1 (ARS (Bidirectionally acting factor)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     55807 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     50.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; Z19865; CAA79673.1; -.
     51.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                        14; Conservative
                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Trans-acting factor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98
291
360
402
496 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
Matches 10; Conserv
Query Match
Best Local Similarity
Matches 14; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=4911;
                                                                                                                                                                                                                                                                                                                                                                    01-FEB-1994
01-JUN-1994
                                                                                                                                                                                                                                                                                                                                              01-FEB-1994
                                                                                                                                                                                                                                                                                           KLUMA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MOD RES
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FING
```

Ŋ

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                      SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

Graham M.R., Lo R.Y.C.;

Graham M.R., Lo R.Y.C.;

Graham M.R., Lo R.Y.C.;

Graham M.R., Lo R.Y.C.;

CARY OUT HIGH-1996) to the EMBL/GenBank/DDBJ databases.

CARY OUT HIGH-AFFINITY BINDING AND ENERGY DEPENDENT UPTAKE INTO THE PERIPLASMIC SPACE OF SPECIFIC SUBSTRATES. IT COULD ACT TO THE PERIPLASMIC SPACE OF SPECIFIC MEMBRANE TO SPECIFIC ENERGY-REQUIRING PROCESSES IN THE OTTOPLASMIC MEMBRANE, RESULTING IN THE RELASE INTO THE PERIPLASM OF LIGAMDS BOUND BY THESE OUTER MEMBRANE PROTEINS (BY SIMILARITY).

CI-SUBCELLIQUAR LOCATION: ANCHORED TO THE CYTOPLASMIC MEMBRANE VIA ITS N-TERMINAL SIGNAL-LIKE SEQUENCE, SPANS THE PERIPLASM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-98150851, PubMed-9491887;
Wei P., Garber M.E., Fang S.-M., Fischer W.H., Jones K.A.;
"A novel CDK9-associated C-type cyclin interacts directly with HIV-1
Tat and mediates its high-affinity, loop-specific binding to TAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Mannheimia.
NCBL_TaxID=75985;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 46.2%; Score 84.5; DB 1; Length 246; Best Local Similarity 50.0%; Pred. No. 0.0014; Matches 17; Conservative 2; Mismatches 8; Indels 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Transport; Protein transport; Inner membrane; Periplasmic;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 7 CYTOPLASMIC (POTENTIAL).
8 28 SIGNAL-ANGHOR (POTENTIAL).
29 246 PERIPLASMIC (POTENTIAL).
246 AA; 27785 MW; C9582F619FCBA5B5 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  117 HKHKHRHKEKREELQEQEKPKDKPKERPKHHKH 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 НКНКНКСК-----НКНКНКНКНКН 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CCT1 HUMAN STANDARD; PRT; 726 AA. 060563; 060581; 16-0CT-2001 (Rel. 40, Created) 16-0CT-2001 (Rel. 40, Last sequence update) 16-0CT-2001 (Rel. 40, Last annotation update) Cyclin T1 (Cyclin T) (CycT1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE=Breast cancer;
MEDLINE=98167917; PubMed=9499409;
Peng J.-M., Zhu Y., Milton J.T., Price D.H.;
                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY: BELONGS TO THE TONB FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Transmembrane; Signal-anchor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; U62565; AAB09530.1; -.
Pasteurella haemolytica.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          rissum=r-cell lymphoma;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Cell 92:451-462(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CCT1_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EWBL outstation the European Bioinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a alicense agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=22484998; PubMed=12597275;
Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
Kohara M., Matsumoto M., Shimpo S., Teuruoka H., Wada T., Yamada M.,
                                                                                                                                                                                                                                                                                                                                                          "Nucleotide sequences of two hydrogenase-related genes (hypA and hypB) from Bradyrhizobium japonicum, one of which (hypB) encodes an extremely histidine-rich region and guanine nucleotide-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                        Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Bradyrhizobiaceae; Bradyrhizobium.
NCBI_TaxID=375;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              46.4%; Score 85; DB 1; Length 302;
46.4%; Pred. No. 0.0015;
tive 1; Mismatches 14; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A -> T (IN REF. 1).
; D3B5F54F24AB90AA CRC64;
                                        01-NOV-1997 (Rel. 35, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
115-SEP-2003 (Rel. 42, Last annotation update)
147PB OR BLL6931.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Last sequence update)
Last annotation update)
302 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Metal-binding; Nickel; Complete proteome. boMAIN 16 54 HIS-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Biochim. Biophys. Acta 1184:135-138(1994)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 КНКНКНКНКСКНКНКНКНКНКНКНКН 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15 ЕНАНОННИРИСКОРНОНСКИНИНИЯ 42
PRT;
                                                                                                                                                                                                                                                                                            STRAIN-USDA 110;
MEDLINE-94137733; PubMed-8305450;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TONB PASHA STANDARD; PR. 1972.04; 1972.04; 15-DEC-1998 (Rel. 37, Created) 15-DEC-1998 (Rel. 37, Last seque: 16-OCT-2001 (Rel. 40, Last annor. TONB protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR004392; HypB.
InterPro; IPR002894; HypB. UreG.
Pfam; PF01495; HypB_UreG; I.
TIGRFAMs; TIGR00073; hypB; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; L24513; AAA17763.1; -.
EMBL; AP005960; BAC52196.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      302 AA; 32708 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 46.4%;
Matches 13; Conservative
                                                                                                                                                          Bradyrhizobium japonicum.
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                         Fu C., Maier R.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-USDA 110
HYPB BRAJA
                                                                                                                                                                                                                                                                                                                                                                                                                                      domains.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
```

TONB_PASHA RESULT 8

884448

ઠે

셤

ï

Gaps

7;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOTALLA TO THE CONTRICT OF THE CYCLIN-DEPENDENT KINASE PAIR

J. Virol. 73:5777-5786(1999).

L. FUNCTION: REGULATORY SUBUNIT OF THE CYCLIN-DEPENDENT KINASE PAIR

CLOK9/CYCLIN T) COMPLEX, ALSO CALLED POSITIVE TRANSCRIPTION

CLOK9/CYCLIN T) COMPLEX, ALSO CALLED POSITIVE TRANSCRIPTION

CLOK9/CYCLIN T) COMPLEX, ALSO CALLED POSITIVE TRANSCRIPTION

CLOK9/CYCLIN T) COMPLEX, ALSO CALLED PRODUCTION ELONGATION BY

TRANSCRITION FROM ABORTIVE TO PRODUCTION ELONGATION BY

TRANSCRIPTION PROMAIN OF THE HIV-1 NUCLEAR TRANSCRIPTIONAL

ACTIVATOR, TAY. IT IS PROBABLY THE CIS-ACTING TRANSCRIPTIONAL

ACTIVATOR, TAY. IT IS PROBABLY THE CIS-ACTING TRANSCRIPTIONAL

ACTIVATOR, TAY. IT IS PROBABLY THE CIS-ACTING TRANSCRIPTIONAL

ACTIVATOR, TAY. IT AS PROBABLY THE CIS-ACTING TRANSCRIPTIONAL

ACTIVATOR, TAY. IT AS PROBABLY THE CIS-ACTING TRANSCRIPTIONAL

ACTIVATOR ASSOCIATES WITH CDK9 TO FORM P-TEPB. CYCLIN TI IS THE

C. ISUBUNIT: ASSOCIATES WITH CDK9.

C. ISUBUNIT: ASSOCIATES WITH CDK9.

C. ISUBUNIT: ASSOCIATES WITH CDK9.

C. ISUBUNIT INTERNATION: NOCIER.

C. ISUBURIT INTERNATION: NOCIER.

C. ISUBURIT INTERNATION: NOCIER.

C. ISUBURIT INTERNATION: NOCIER.

C. INSCRILLANEOUS: INTERNACTION BETWEEN TAT AND CYCLIN TI REQUIRES
                                                                                                  MEDLINE-99263518; PubMed-10329125;
Ivanov D., Kwak Y.T., Nee E., Guo J., Garcia-Martinez L.F.,
Gaynor R.B.;
"Cyclin II domains involved in complex formation with Tat and TAR RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-99051315; PubMed-8832504;
Garber M.E., Wei P., KewalRamani V.N., Mayall T.P., Herrmann C.H.,
Rice A.P., Littman D.R., Jones K.A.;
"The interaction between HIV-1 Tat and human cyclin T1 requires zinc
"The distraction between tivel tat and human cyclin T1 requires zinc
"T4 a critical cysteine residue that is not conserved in the murine
Cycll protein.;
Genes Dev. 12:3512-3527(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FUNCTION.
MEDLINE-992873; PubMed=10364129;
Bieniasz P.D., Grdina T.A., Bogerd H.P., Cullen B.R.;
Bieniasz P.D. Grdina T.A., Bogerd H.P., Cullen B.R.;
"Analysis of the effect of natural sequence variation in Tat and in cyclin T on the formation and RNA binding properties of Tat-cyclin T
                                                                                                                                                                                                                                                                                     MEDLINE=99263519; PubMed=10329126;
Kwak Y.T., Ivanov D., Guo J., Nee E., Gaynor R.B.;
"Role of the human and murine cyclin T proteins in regulating HIV-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -i- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN C SUBFAMILY.
"Identification of multiple cyclin subunits of human P-TEFD."; Genes Dev. 12:755-762(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MIM; 602506; -
603 GO:00000019; P:regulation of CDK activity; TAS.
GO; GO:000016; P:transcription from Pol II promoter; TAS.
InterPro; IPR006670; Cyclin.
InterPro; IPR006671; Cyclin.N.
                                                                                                                                                                                                  are critical for tat-activation.";
J. Mol. Biol. 288:41-56(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; AF048161; AAC39638.1; -. EMBL; AF048730; AAC39664.1; -.
                                                                                                                                                                                                                                                                                                                                                                                          Mol. Biol. 288:57-69(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00134; cyclin; 1.
SMART; SM00385; CYCLIN; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSFAC; T02727; -.
Genew; HGNC:1599; CCNT1.
GK; O60563; -.
                                                                                                                                                                                                                                                                                                                                                                      Tat-activation.";
                                                                                                                                                                                                                                                                       MUTAGENESIS
```

```
MEDLINE-99303751; PubMed=10373508;

MEDLINE-99303751; PubMed=10373508;

MEDLINE-99303751; PubMed=10373508;

MEDLINE-99303751; PubMed=10373508;

MIGHALY divergent lentiviral Tat proteins activate viral gene rexpression by a common mechanism.";

Texpession by a common mechanism.";

MOI. Cell. Biol. 19:4592-4599(1999)

CCDK9/CYCLIN T) COMPLEX, ALSO CALLED POSITIVE TRANSCRIPTION ELONGATION FACTOR B (P-TEFB), WHICH IS PROPOSED TO FACILITATE THE ELONGATION PACTOR B (P-TEFB), WHICH IS PROPOSED TO FACILITATE THE ELONGATION PEROPORTION BY THE BOUNT OF THE LARGE SUBDINT OF RNA POLYMERASE II (RNAP II) (BY SIMILARITY). BINDS ALSO COPACTOR FOR TAT. ODES NOT BIND TO THE TRANSCRIPTIONAL ACTIVATOR, TAT. IT IS PROBABLY THE CIS-ACTING TRANSCRIPTIONAL ACTIVATOR. TAT. IT IS DOMAIN OF THE BUILD SANDAR ACTIVATION PROBABLY THE LISTANSACTIVATION COPACTOR FOR TAT. DOES NOT BIND TO THE TRANSACTIVATION DOMAIN OF THE BUILD SANDAR ACTIVATOR.

COMMON THE HIV-1 NUCLEAR TRANSCRIPTIONAL ACTIVATOR.

COMMON THE HIV-1 NUCLEAR TRANSCRIPTIONAL ACTIVATOR.

COMMON THE HIV-1 NUCLEAR TRANSCRIPTIONAL ACTIVATOR.

CHI-SUBDING COPACTOR FOR TAT. DOES NOT BIND TO THE TRANSACTIVATION DOMAIN OF CYCLIN ASSOCIATED WITH CDKS TO CHANTAL ACTIVATOR.

CHI-SUBDING COPACTOR TRANSCRIPTIONAL ACTIVATOR.

CHI-SUBDING TO THE TRANSCRIPTIONAL ACTIVATOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                    2
                                                (POTENTIAL).
                                                                                                         POLY-HIS.
POLY-SER.
POLY-SER.
C-Y: LOSS OF HIV-1 TAT TRANSACTIVATION.
Q -> R (IN REF. 2).
4637EFB2DDEDFE13 CRC64;
                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Būkaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
                                                                                                                                                                                                                                                                                      4
          Length 726;
                                                                                                                                                                                                                                                         0.0045;
                                                                                                                                                                                                                                                   DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16-OCT-2001 (Rel. 40, Last sequence update) 16-OCT-2001 (Rel. 40, Last annotation update) Cyclin T1 (Cyclin T) (CycT1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  727 AA
                                                                                                                                                                                                                                                                                        1; Mismatches
                                                                                                                                                                                                                                                                                                                                                           505 EHKEKHKTHPSNHHHHHNH---HSHKHSH 530
                                                                                                                                                                                                                                                                                                                              1 КНКНКНК-НКСКНКНКНКНККККНКНКНКН 28
                                                                                                                                                                                                                                                   Score 84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfan; PF00134; CyClin; 1.
SMART; SM00385; CYCLIN; 1.
PROSITE; PS00292; CYCLINS; FALSE_NEG.
                                                                                                                                                                                                                                                                       Pred. No.
PROSITE; PS00292; CYCLINS; FALSE_NEG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; AF137509; AAD38518.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR006670; Cyclin.
InterPro; IPR006671; Cyclin.N.
                                                                                                                                                                                                                 80684 MW;
                                                                                                                                                                                                                                                 45.98;
                                                                                                                                                                                                                                                                       51.7%;
                                                                                                                                                                                                                                                                                          15; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Equus caballus (Horse)
                  Cyclin; Cell cycle;
                                                                                                                                                                                                                 726 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. TISSUE=Fibroblast;
                                                                                                                                                                                                                                                                       Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=9796;
                                      Transcription
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HORSE
                                                                                                                                                                        MUTAGEN
                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9XTZ6;
                                                                           DOMAIN
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                    CCT1_HORSE
                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DR
KW
KW
KW
FT
FT
FT
FT
SO
                                                                                                                                                                                                                                                                                                                                    ò
```

```
M60703;
  sclerosis.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL;
EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL;
~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 657-765 FROM N.A.
MEDLINE=90046823; PubMed=2479024;
Maul G.G., Jimenez S.A., Riggs E., Ziemnicka-Kotula D.;
"Determination of an epitope of the diffuse systemic sclerosis marker antigen DNA topoisomerase I: sequence similarity with retroviral p30gag protein suggests a possible cause for autoimmunity in systemic
                regulation; Nuclear protein.
3 270 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
4 272 TAT:TAR RECCENTION MOTIF (TRM) (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Fujimori A., Harker W.G., Konlhagen G., Hoki Y., Pommier Y.; "Mutation at the catalytic site of topoisomerase I in CEM/C2, a human leukemia cell line resistant to camptothecin.";
                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Kunze N., Yang G., Dolberg M., Sundarp R., Knippers R., Richter A., "Structure of the human type I DNA topoisomerase gene."; J. Biol. Chem. 266:9610-9616(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              D'ATDA P., Machlin P.S., Ratrie H. III, Rothfield N.F.,
Cleveland D.W., Earnshaw W.C.;
Cleveland D.W., Earnshaw W.C.;
"CDNA cloning of human DNA topoisomerase I: catalytic activity of
67.7-ba cataboxyl-terminal fragment.";
Proc. Natl. Acad. Sci. U.S.A. 85:2543-2547(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 541-765 FROM N.A.
MEDLINE=89288043; PubMed=2544263;
Zhou B.S., Bastow K.F., Cheng Y.C.;
"Characterization of the 3' region of the human DNA topoisomerase
                                                                                                                                                                                                                                                       4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 5-765 FROM N.A., AND VARIANTS THR-370 AND SER-722 IISSUE=Peripheral blood; MEDLINE=95188170; Pubmed=7882333;
                                                                                                                                                                                                               Length 727;
                                                                                                                                                                                                         Score 84; DB 1; Length 727
Pred. No. 0.0045;
1; Mismatches 9; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                       BFC2A398D6B35BCE CRC64;
                                                                    COLLED COIL (POTENTIAL)
POLY-HIS.
division; Coiled coil;
                                                                                                                                                                                                                                                                                                                                                                                                                      TOP1 HUMAN STANDARD; PRT; 765 AA. P11387; 041256; 012855; 012856; 090JNO; 01-JUL-1989 (Rel. 11, Created) 116-OCT-2001 (Rel. 40, Last sequence update) 15-SEP-2003 (Rel. 42, Last annotation update) DNA topoisomerase I (EC 5.99.1.2).
                                                                                                                                                                                                                                                                                                                              507 EHKEKHKTHPSNHHHHHNH---HSHKHSH 532
                                                                                                                                                                                                                                                                                         1 КНКНКНК-НКСКНКНКНКНКНКНКН 28
                                                                                                                                  POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=91236733; PubMed=1851751;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=88190108; Pubmed=2833744;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Cancer Res. 49:3922-3927 (1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cancer Res. 55:1339-1346(1995)
                                                                                                                                                                         81013 MW;
                                                                                                                                                                                                             45.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 1-436 PROM N.A.
                                                                                                                                                                                                                               Best Local Similarity 51.7
Matches 15; Conservative
Cell
                                                                                                              528
573
726
                                                                                          386 4
519 5
562 5
718 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens (Human)
Cyclin, Cell cycle;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                Franscription
DOMAIN 25
                                                                                                                                                  DOMAIN
SEQUENCE
                                                                                                                                                                                                           Query Match
                                                      DOMAIN
                                                                                            DOMAIN
                                                                                                                                  DOMAIN
                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                          TOP1 HUMAN
                                                                                                                                                                                                                                                                                                                                                                                        RESULT 11
SHIFFFF
                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOTTO 
                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                              g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SUBUNIT: MONOMER.
MISCELLANEOUS: EUVRAYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
RELAX ONLY NEGATIVE SUPERCOILS.
MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA
BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN
WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS
AT ONE END OF THE ENZYME-SEVERED DNA STRAND.
SIMILARITY: BELONGS TO THE EUKARYOTIC TYPE I TOPOISOMERASE FAMILY.
                                                                                                                                                                                                                                                              VARIANT CPT-RESISTANT ALA-729.
MEDLINE=93075133; PubMed=1332703;
Kubota N., Kanzawa F., Nishio K., Takeda Y., Ohmori T., Fujiwara Y., Terashima Y., Saijo N.;
"Detection of topoisomerase I gene point mutation in CPT-11 resistant lung cancer cell line.";
Biochem. Biophys. Res. Commun. 188:571-577(1992).
                                                 VARIANTS CPT-RESISTANT.

MEDIATNE-91187651; PubMed=1849260;

MEDIATNE-91187651; PubMed=1849260;

Keene J.D., Okada K., Kjeldsen E., Nishikawa K.;

"Molecular cloning of a cDNA of a camptothecin-resistant human DNA topolsomerses I and identification of mutation sites.";

Nucleic Acids Res. 19:69-75(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 215-765.
MEDLINE=98155246; PubMed=9488644;
Redinbo M.R., Stewart L., Kuhn P., Champoux J.J., Hol W.G.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Crystal structures of human topoisomerase I in covalent and noncovalent complexes with DNA."; Science 279:1504-1513 (1998).
Proc. Natl. Acad. Sci. U.S.A. 86:8492-8496(1989)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAA61206.1; JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; J03250; AAA61207.1; -.
EMBL; M60706; AAA61206.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAA61206.1;
AAA61206.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAA61206.1;
AAA61206.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAA61206.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAA61206.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAA61206.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAA61206.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAA61206.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA61206.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA61206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAA61206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAA61206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA61206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              M60688;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            M60689;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           M60691;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              M60699;
```

```
29 K 29
                                                                             82
                                                                                                                                                                Cricetulus.
                                                                             -14
                                                                                                         CRIGR
                      Query Match
Best Local S:
Matches 22;
                                                                             82
HELIX
TURN
TURN
                                                                                                              셤
                                                                   ð
                                                                                 GK; Pill387; --

MM, 126420; --

GO, GOO:0003917; F:DNA topoisomerase I activity; TAS.

InterPro; IPR001631; Topismerse I.

InterPro; IPR001631; Topismerse I.

Pfam; PF002919; Topoisomerase I; I.

Pfam; PF001028; Topoisomerase I; I.

PRINTS; PR00416; EUTPISMRASEI.

PRNATT; SM00435; TOPEUc; I.

PROSTITE; PS00176; TOPOISOMERASE I EUK; I.

SNORTI; PR00176; TOPOISOMERASE I EUK; I.

ISOMERSASE; DNA-DINGING; POlymorphism; 3D-structure.
                                                                                                                                         DNA CLEANGE (BY SIMILARITY).
M -> T (IN CPT-RESISTANT CELL LINE CEMC2).
/FTId=VAR 010666.
D -> G (IN CPT-RESISTANT CELL).
/FTId=VAR 007530.
N -> S (IN CPT-RESISTANT CELL LINE CEM/C2).
                                                                                                                                                                                        /FTIG=VAR_010667.
T -> A (TN CPT-RESISTANT CELL).
/FTIG=VAR_007531.
A -> V (IN REF. 1).
 EMBL; M60704; AAA61206.1; JOINED.
EMBL; M60705; AAA61206.1; JOINED.
EMBL; U07804; AAB60399.1; --
EMBL; U07806; AAB60380.1; --
EMBL; AL038652; CAB43980.1; --
EMBL; M27913; AAA61208.1; --
EMBL; M31913; AAA61208.1; --
PIR; A30897; ISHUTI.
PDB; 1A31; 19-AUG-98.
PDB; 1A36; 12-AUG-98.
PDB; 1A36; 12-AUG-98.
PDB; 1B35; 26-AUG-98.
PDB; 1B35; 26-AUG-98.
PDB; 1B36; 12-AUG-98.
PDB; 1B36; 12-AUG-98.
PDB; 1B36; 12-AUG-98.
PDB; 1B36; 12-AUG-00.
Aarhus/Ghent-2DPAGE; 610; NEPHGE.
                                                                                                                                        POLY-LYS
                                                                                                                                             723
                                                                                                                                                                                                                                                 533
                                                                                                                                                                               722
                                                                                                                                        DOMAIN
ACT SITE
VARIANT
                                                                                                                                                                                                           CONFLICT
STRAND
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                               STRAND
TURN
STRAND
STRAND
                                                                                                                                                                                                                                                                 VARIANT
                                                                                                                                                                                                VARIANT
                                                                                                                                                                                                                                      TURN
                                                                                                                                                                                VARIANT
                                                                                                                                                                                                                                 TRAND
                                                                                                                                                                                                                                                                                                                                                                                                      TURN
HELIX
TURN
                                                                                                                                                                                                                                                  TELIX
```

```
DROME
                                                                   ACT SITE
CONFLICT
                                                                                             CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                          CONFLICT
                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                      CONFLICT
                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                Local
                                                                                                                                                                                                                                                                            Matches
    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hui C.F., Lo C.K., Hwang J.;

Hui C.F., Lo C.K., Hwang J.;

SEQUENCE FROM N.A.

Hui C.F., Lo C.K., Hwang J.;

SUBMITTED BY TOPOISOMERASES LEADS TO THE SUBMITTED BY TOPOISOMERASES LEADS TO THE CONVERSION OF ONE TOPOLGGICAL ISOMER OF DNA TO ANOTHER.

- I- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded DNA, followed by passage and rejoining.

- I- SUBUNIT: Monomer.

- SUBUNIT: Monomer.

- WISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH NEGALIVE AND POSITIVE SUPERCOILS.

- I- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS COME AT ONE END OF THE ENZYME IS JOINED TO A DNA PHOSPHORUS COME AT ONE END OF THE ENZYME IS TRAND.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                      2 HKHKHKH------KHKHKHKHKGKH------KHKH 26
                                                                                                                                                            HKHKDKHKDREHRHKEHKKDKEKDREKSKHSNSEHKDSEKKHKEKEKTKHKDGSSEKHKD 81
                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus.
                                                                                                            35;
PROSITE; PS00176; TOPOISOMERASE I EUK; 1.
ISOMERASE; TOPOISOMERASE I DNA LEAVAGE (BY SIMILARITY).
ACT SITE 725 DNA CLEAVAGE (BY SIMILARITY).
VARTANT 505 505 S -> G (IN CPT-RESISTANT CELL).
SEQUENCE 767 AA; 90867 MW; 347336D424EF35A9 CRC64;
                                                                               Score 82.5; DB 1; Length 767;
Pred. No. 0.007;
0; Mismatches 6; Indels 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
MEDLINE=93146125; PubMed=8096488;
KOIWAI O., Yasui Y., Sakai Y., Watanabe T., Ishii K.,
Yanagihara S., Andoh T.;
"Cloning of the mouse cDNA encoding DNA topoisomerase I and
cencensomal location of the gene.";
Gene 125:211-216(1993).
                                                                                                                                                                                                                                                                                                               01-0CT-1993 (Rel. 27, Created)
01-0CT-1993 (Rel. 27, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
TOP1 OR TOP-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001631; Topismerse
Pfam; PF02919; Topoisomer_I_N; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; D10061; BAA00950.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       L20632; AAA40466.1; -.
                                                                                             Local Similarity 34.9%;
nes 22; Conservative
                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AGD; MGI:98788; Topl.
                                                                                                                                                                                                                                                                                                                                                                                     Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HSSP; P11387; 1A35.
                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=10090;
                                                                                                                                                                                          27 KHK 29
                                                                                                                                                                                                                     KHK 84
                                                                                                                                                                                                                                                                                     TOP1 MOUSE
Q04750;
                                                                                 Query Match
                                                                                                                                                                                                                                                           RESULT 13
TOP1 MOUSE
                                                                                               Best Loca
Matches
                                                                                                                                      8
                                                                                                                                                                                           δ
                                                                                                                                                                                                                   윱
```

```
A Adams M.D., Celniker S.B., Holf R.A., Evans C.A., Gocayne J.D.,
A Adams M.D., Celniker S.B., Holf R.A., Evans C.A., Gocayne J.D.,
A Adams M.D., Celniker S.B., Holf R.A., Hoskins R.A., Galle R.F.,
A George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
A Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
A Marn K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
A Abril J.F., Agbayani A., An H.-J., Bayraktaroglu L., Beasley B.M.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Berson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolahakov S.,
Borkova D., Botchan M.R., Bouck J., Borstein P., Brottier P.,
Burtis K.C., Busam D.A., Buller H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Gawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
A Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Pleischmann W.,
A Durbin K.J., Evangelista C.C., Ferraz C., Ferriers S., Fleischmann W.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----KGKH------KHKHKHKGKH-----KHKH 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22 HKHKDKHKDREHRHKEHKKDKDKDREKSKASNSEHKDSEKKHKEKEKTKHKDGSSEKHKD 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Verrijzer C.P., Yokomori K., Chen J.-L., Tjian R.;
"Drogophila TAFII150: similarity to yeast gene TSM-1 and specific
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-FEB-2003 (Rel. 41, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Transcription initiation factor TFIID 150 kDa subunit (TAFII-150)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Drosophila melanogaster (Fruit fly).

Bukaryota, Metazoa, Arthropoda; Hexapoda, Insecta; Pterygota; Neoptera; Brachycera; Muscomorpha; Ephydroidea, Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 82.5; DB 1; Length 767;
Pred. No. 0.007;
0; Mismatches 6; Indels 3!
                                                                                                                                                    SIMILARITY)
                                                                                                                                                                       R -> P (IN REF. 2).
D -> E (IN REF. 2).
A -> V (IN REF. 2).
MISSING (IN REF. 2).
S -> L (IN REF. 2).
G -> B (IN REF. 2).
G -> E (IN REF. 2).
G -> V (IN REF. 2).
G -> V (IN REF. 2).
D -> Y (IN REF. 2).
M; 398327062B179F2A CRC64;
Pfam; PF01028; Topoisomerase I; 1.
PRINTS; PR00416; EUTPISNRASEI.
SMART; SM00435; TOPEUC; 1.
PROSTIE; PS00176; TOPOISOMERASE I EUK; 1.
Isomerase; Topoisomerase; DNA-binding.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=20196006; PubMed=10731132;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE=Embryo;
MEDLINE=94233377; PubMed=8178153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     binding to core promoter DNA.";
Science 264:933-941(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   90789 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      024325; Q8SZR7; Q9VT64;
01-NOV-1997 (Rel. 35, Created)
28-FEB-2003 (Rel. 41, Last seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        45.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TAF2 OR TAF150 OR CG6711.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2 HKHKHKH------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22; Conservative
                                                                                                                                                       725
91
1121
1121
1121
161
167
277
277
292
533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                767 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ||||
KHK 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN=Berkeley;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       27 KHK 29
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=96433160; PubMed=8836188;

AEDLINE=96433160; PubMed=8836188;

AEDLINE=96433160; PubMed=8836188;

AEDLINE=96433160; PubMed=8836188;

AEDLINE=96433160; PubMed=8836188;

AEDLINE=96433160; PubMed=8836188;

AEDLINE=12 Cloud of the gene for the somatic form of DNA copoieomeraee I from Xenopus laevis.";

RI TOPOISOMERASES LEADS TO THE TOPOISOMERASES LEADS TO THE CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.

CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.

CONVERSION OF ONE TOPOLOGICAL ISOMER OF SINGLE-STRANGE OF SINGLE-STRANGE
                                                                                                                                                                                                                                                                           ı,
                                                                                                  DKOKERKOKKOKROKHISKLQARETATPOTLSSEDSSNSNSL
PPWNIN -> ERKOKOKROPHISRLQAARQPLRTLSARRTV
ATAIACRP (IN REF. 1).
                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
Xenopodinae; Xenopus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               R MBL; L07777; AAB36608.1; -.
R PIR; S72366; S72366
R HSSP; P11387; 1A35.
R InterPro; IPR001631; Topismerse_I.
R Pf02919; Topoisomer I N; 1.
R Pf02919; Topoisomer I N; 1.
R Pf02919; Topoisomer I N; 1.
R PRINTS; PR00416; EUTPISMRASEI.
R SMART; SM00435; TOPEUC; 1.
R PROSITE; PS00176; TOPOISOWERASE I EUK; 1.
R PROSITE; PS00176; TOPOISOWERASE I EUK; 1.
R ISOMERASE; TOPOISOWERASE; DNA-binding.
R ACT_SITE ; 779 779 DNA CLEAVAGE (BY SIMILARITY).
                                                                                                                                                                                                                                                                           1;
                                                                                                                                                                                                                               DB 1; Length 1221;
                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                   MW; C2DC066826B1AF6E CRC64;
lear protein.
BINDS TO TBB AND TAFII-250.
HIGHLY CHARGED.
R -> S (IN REF. 1).
H -> P (IN REF. 1).
                                                                                                                                                                                                                               Score 81.5; DB 1
Pred. No. 0.014;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                       829 AA
                                                                                                                                                                                                                                                                                2; Mismatches
                                                                                                                                                                                                                                                                                                                          1 КНКНКНКНКБКНКНКНКН-КБКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Xenopus laevis (African clawed frog)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DNA topoisomerase I (EC 5.99.1.2).
    regulation; Nuclear
                                                                                                                                                                                                                                      44.5%;
56.7%;
                                                                                                                                                                                     1221 AA; 139498
                                                                                                                                                                                                                                                            Local Similarity 56.7 tes 17; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                    53
88
1221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=8355;
                             845
1138
53
88
1175
            Transcription
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     XENLA
                                                                         CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOP1 XE
                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 15
TOP1_XENLA
                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NA FIT FIT FIT SO SET SO
                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                셤
```

```
Query Match 44.0%; Score 80.5; DB 1; Length 829; Best Local Similarity 41.9%; Pred. No. 0.013; Matches 18; Conservative 2; Mismatches 8; Indels 15; Gaps
                                                                                                                         SQ SEQUENCE 829 AA; 98230 MW; 8D1FE4252A916219 CRC64;
                                                                                                                              ઠે
```

5;

Search completed: January 20, 2004, 18:23:54 Job time : 10.6667 secs

g

	The second of th			And the second of the second o	•
					:
		* -	•		
					•
				•	
			•		
		at .			•
		×.			
		*		•.	
		*			
	•				
					10
		* :			
		·			
			•	•	e Pr
	The second secon				i i g
					1
					W.
				*	
	•				
					4
	•				
	% V	* *			
			•		
				May and the same of the same o	
		*			
		* * * *			
곳() (1985년 - 1984년 - 1987년 - 1			4		
		4 4 4			
				1	· · · · · · · · · · · · · · · · · · ·
			4 ·		
	- y g v , · · ·		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		

Wed

Perfect score:

Sequence:

OM protein

Run on:

Scoring table:

Searched:

Database

```
Q9zw23 arabidopsis
Q9v714 drosophila
Q8i418 plasmodium
Q8swg4 encephalito
Q8stf5 dictyosteli
Q8ipf9 drosophila
Q9i488 plasmodium
Q9nw94 homo sapien
Q9brr8 homo sapien
Q8izv6 plasmodium
Q8izv6 plasmodium
Q8izv7 homo sapien
Q8i399 plasmodium
Q8ico4 plasmodium Q8i3v1 plasmodium Q8i2u9 plasmodium Q8ikj8 plasmodium Q8ikj8 plasmodium Q9vhp7 drosophila Q9xe55 ralstonia s
                                                                                                                                                                                                                                                                                                                                                                                                                                        Q8x3u5 escherichia
Q8t0g9 drosophila
Q8hxn7 pan troglod
                                                                                                                                                                                                                                                                                                                                                                                                                           Q8kbi9 chlorobium
                                                                                                                                                                                                                                                                                                                                                                     Q8ie91 plasmodium
Q8if19 plasmodium
Q8i417 plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-22255705; PubMed=12368864;

MEDLINE-22255705; PubMed=12368864;

MEDLINE-22255705; PubMed=12368864;

Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,

Carlton J.M., Pain A., Nelaon K.B., Bowman S., Paulsen I.T., James K.,

Bisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,

Chan M.-S., Nene V., Shallom S.J., Bub B., Peterson J., Angluoli S.,

Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,

Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,

McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,

Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,

"Genome sequence of the human malaria parasite Plasmodium

falciparum.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Plasmodium falciparum (isolate 3D7).
Eukaryota, Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=36329;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 127; DB 5; Length 954;
Pred. No. 1.4e-08;
8; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nature 419:498-511(2002).
EMBL, AE014818, AAN36747.1; -.
Hypothatical protein.
SEQUENCE 954 AA, 114907 MW; 8BD6834A5846F21C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ALIGNMENTS
 081C04
0812U9
081ZU9
081KJ8
098KJ8
092K23
092K23
092K23
091T44
081FF5
081FF5
081FF9
0911V0
081E91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ...cmsE-2003 (TrEMBLrel. 23, Ca. 01-MAR-2003 (TrEMBLrel. 23, La. Hypothetical protein. PP14_0135.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 69.4%;
Best Local Similarity 62.1%;
Matches 18; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
     SEQUENCE FROM N.A.
     554.1
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
550.8
500.8
500.8
500.8
500.8
500.8
500.8
500.8
500.8
500.8
500.8
500.8
                                                                                                                                                                                                                    Q81LV8;
01-MAR-2003
                                                                                                                                    RESULT 1
     QBILVB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q8ic24 plasmodium o68651 haemophilus O66651 mus musculu Q9ngh2 homo sapien Q8hnd8 mus musculu O15793 plasmodium Q8st04 dictyosteli o43604 homo sapien O60668 homo sapien O43487 homo sapien O43487 homo sapien Q8suk4 encephalito Q8suk4 encephalito Q8tiw6 methanosarc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   O8ilv8 plasmodium
O8ic24 plasmodium
O68651 haemophilus
                                                                                                                                       (without alignments)
209.478 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                      January 20, 2004, 18:20:39; Search time 35.7246 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Description
                    GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                    Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                      183
1 КНКНКНКНКСКНКНКНКНККККНКНК 29
                                                                                                                                                                                                                                                                                                         830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                     - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sp_rodent:*
sp_virus:*
sp_vertebrate:*
sp_unclassified:*
sp_rvirus:*
sp_rvirus:*
                                                                                                                                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_human:*
sp_invertebrate:*
sp_mammal:*
sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_organelle:*
sp_phage:*
sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_archea:*
sp_bacteria:*
sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                         Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  archeap:*
                                                                                                                                                                                        US-10-018-103A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SPIREMBL 23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1461
11189
11199
11199
260
260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1007
1007
776
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9:
```

ö

Gaps

ö

1 КНКНКНКНКСКНКНКНКНКНКНК 29

g ઠ

2

1110.5 1110.5 1110.5 110.5 104.5 104.5 104.5 104.5

Regult No.

366 AA.

PRT;

Created)

```
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Pasteurella.
NCBI_TaxID=747;
                                                  01-JUN-2001 (TrEMBLrel. 17, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                      STRAIN=Pm70;
MEDLINE=21145866; PubMed=11248100;
                                         01-JUN-2001 (TrEMBLrel. 17,
                   PRELIMINARY;
                                                                                       FIMA OR PM0926.
Pasteurella multocida.
                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=10090;
                                 Q9CMAB;
                     Q9CMA8
                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 5
                                                                              FimA.
RESULT 4
            Q9CMA8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Lu D., Boyd B., Lingwood C.A.; ^{\prime} "The expression and characterization of a putative adhesin B from H. influenzae.";
                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN=6564;
MEDILINE=98030581; PubMed=9360976;
MEDILINE BOY B., Lingwood C.A.;
In D., Boyd B., Lingwood C.A.;
"Identification of the key protein for zinc uptake in Hemophilus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                ż
                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Haemophilus.
                                                                                                                                                                                                                                         Score 119; DB 5; Length 3949;
Pred. No. 6e-07;
9; Mismatches 3; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; Score 114; DB 2; Length 311; Pred. No. 2.2e-07; 3; Mismatches 6; Indels
                                                                                                                                                       SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
Seeger K., Murphy L., Harris D., Berriman M., Pain A., Hall
Quail M., Barrell B.;
Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
BMBL, AL844506; CAD50817.1; -.
Hypothetical protein.
SEQUENCE 3949 AA; 467553 MW; A048A9B46EBBEEIB CRC64;
                                         QBIC24;
QBIC24;
QBIC24;
QBIC24;
QBIC24;
QBIC2403 (TrEMBLrel. 23, Last sequence update)
QBI-MAR-2003 (TrEMBLrel. 23, Last annotation update)
QBI-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Hypothetical protein.
MAL7P1.19
Plasmodium falciparum (isolate 3D7).
Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      311 AA; 34957 MW; 6F0F1D559C867E8B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                           07, Last sequence update)
22, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                            311 AA
                             PRT; 3949 AA
                                                                                                                                                                                                                                                                                                        1567 EHEHEHEHEHEHEHEHEHEHEHEHEHEHEN 1594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEMS Microbiol, Lett. 165:129-137(1998).
EMBL, AF053125; AAC08991.1; -.
InterPro, IPR006127; SBP_bac_9.
Pfam; PF01297; SBP_bac_9; 1.
                                                                                                                                                                                                                                                                                          1 КНКНКНКНКСКНКНКНКНКНКНКНКН 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 НХНКНКСКНКНКНКНКСКНКНКК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 influenzae.";
J. Biol. Chem. 272:29033-29038(1997).
                                                                                                                                                                                                                                                                                                                                                                                         O68651;
01-AUG-1998 (TrEMBLrel. 07, Created)
                                                                                                                                                                                                                                                                                                                                                                                                               01-AUG-1998 (TrEMBLrel. 07, Last sec 01-OCT-2002 (TrEMBLrel. 22, Last am Putative adhesin B FimA (Fragment). Haemophilus influenzae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62.3%;
                                                                                                                                                                                                                                             Query Match
Best Local Similarity 57.1%;
Matches 16; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
Matches 17; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=6564;
MEDLINE=98377728;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=727;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TER
                                                                                                                                                                                                                                                                                                                                                                               068651
                               Q81C24
                                                                                                                                                                                                                                                                                                                                                       RESULT 3
068651
         RESULT 2
                                                                                                                                                                                                                                                                                                                                                                               ò
                                            ò
                                                                                                                                                                                                                                                                                                                      셤
```

```
э
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=C57BL/6J; TISSUE=Sympathetic ganglion;
MEDLINE=22354683; PubMed=12466851;
The FANTOW Consortium,
the RIKEN Genome Exploration Research Group Phase I & II Team;
than Laysis of the mouse transcriptome based on functional annotation of the mouse transcriptome based on function of the mouse transcriptome based on function
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gарв
May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
"Complete ganomic sequence of Paustian M.L., Whittam T.S., Kapur V.;
Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).

EMBL, AENGA12; ARK03101.;
InterPro; IPR006129; Kininogen.
InterPro; IPR0061295; Kininogen.
InterPro; IPR0061295; Kininogen.
Pfam, PF01277; SBP bac. 9.
Pfam, PF01277; SBP bac. 9.
Pfam, PF01277; SBP bac. 9.
Prants; PR00590; ADHESNPAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ch 60.4%; Score 110.5; DB 11; Length 497;
1 Similarity 59.0%; Pred. No. 9.9e-07; 23; Conservative 1; Mismatches 4; Indels 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 366;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    366 AA; 41443 MW; BF6A5929B78EC13D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-MAR-2003 (TrEMBLrel. 23, Created)
01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 КН----КНКНКСКСНКНКНКС-----КНКНКНКК
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61.7%; Score 113; DB 16;
63.0%; Pred. No. 3.5e-07;
tive 2; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   497 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               159 HXHDHDHKHDHAHKHEHDHKHDHEHKH 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 нхнкнкнкнкнкнкнкнкнкн 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pre-mRNA protein kinase (Fragment).
Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
Matches 23; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Complete proteome.
SEQUENCE 366 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
```

셤

1007 AA

```
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 . Query Match
Best Local Similarity
                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                      NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (cg7) genea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN=Dd2;
                                                                                                                                                                                                                                                                                                                                                                                             015793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                   RESULT 9
g
                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                   ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3;
                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                       InterPro; IPR000719; Prot kinase.
InterPro; IPR002290; Ser thr_pkinase.
InterPro; IPR002290; Ser thr_pkinase.
Probom; PD000001; Prot kinase; 1.
SMART; SM00220; S TKC; 1.
SMART; SM00220; S TKC; 1.
PROSITE; PS500111; PROTEIN KINASE DOM; 1.
PROSITE; PS00108; PROTEIN KINASE ST; 1.
ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
                                                                                                                                       Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
NCBI_TaxID=9606,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                          Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
EMBL; AL033383; CAB94780.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                   11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 60.4%; Score 110.5; DB 4; Length 1007; Best Local Similarity 59.0%; Pred. No. 2e-06; Matches 23; Conservative 1; Mismatches 4; Indels 11;
                                                                                                                                                                                                                                                                                                                                                                                           DB 4; Length 991;
                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Strausberg R.;
Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
EMBL; BC034969; AAH34969.1; -.
SEQUENCE 1007 AA; 116990 MW; 5EEAAFBAEBA7ACB9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                    991 AA; 115247 MW; F809E995B20671D2 CRC64;
                                                                  01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DJ1013A10.1 (PRP4 protein kinase homolog) (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                         1 KH----KHKHKHKHKHKHKHKHKG-----KHKHKHKHK 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAR-2003 (TrEMBLrel. 23, Created)
01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
PRP4 pre-mRNA processing factor 4 homolog B (Yeast)
Homo sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 KH----KHKHKHKGKHKHKHKHKG-----KHKHKHKHK 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       39 KASRHKKKKCHKARSKHK-KHKHSSEEDKDKKHKHKHKHK 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                        60.4%; Score 110.5; DB 4
59.0%; Pred. No. 1.9e-06;
ive 1; Mismatches 4
                                             991 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 1007 AA.
                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity 59.0%;
les 23; Conservative
                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                 DJ1013A10.1.
Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. TISSUE=Brain;
                                                                                                                                                                                                                                               HSSP; Q00534; 1BI7
                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                25
                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                            Q9NQH2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q8IVC3
                                                                                                                                                                                                                                                                                                                                                                                                         Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 7
                                 29NQH2
                                              ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
```

```
ä,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-JAN-1998 (TTEMBLrel. 05, Last sequence update)
01-CT-2002 (TEMBLrel. 22, Last annotation update)
Strain Dd2 heat shock protein 86 (HSP86), 01 (01), 03 (03), 02 (02),
CG8 (CG8), CG4 (Cg4), CG3 (Cg3), uptative chloroquine resistance
transporter (crt), CG9 (cg9), CG1 (cg1), CG6 (cg6), CG2 (cg2), and CG7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Fidock D.A., Dozekunov S.M., Fidock D.A., Nomura T., Talley A.K., Cooper R.A., Dzekunov S.M., Ferdig M.T., Ursos L.M.B., bir Singh Sidhu A., Naude B., Deitsch K.W., Su X.E., Woctron J.C., Roepe P.D., Wellems T.E.; mutations in the P. falciparum digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance."; Mol. Cell 6:861-871(2000).

EMBL, AP030694; AAC47848.1; ... SEQUENCE 776 AA; 92634 MW; 35731749BACCA8EC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        oŧ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                     STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
MEDLINE=22354683; PubMed=12466851;
The FANTOM Consortium,
The RIKEN Genome Exploration Research Group Phase I & II Team;
"Analysis of the mouse transcriptome based on functional annotation 60,770 full-length consa.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Su X., Kirkman L.A., Fujioka H., Wellems T.E.;
"Computer polymorphisms in an approximately KDa protein are linked to colloroquine-resistant P. falciparum in Southeast Asia and Africa.";
Cell. 91:593-603(1997).
                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 110.5; DB 11; Length 1007;
Pred. No. 2e-06;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Plasmodium falciparum (isolate Dd2).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=57267;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 776;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               59.0%; Score 108; DB 5; Length 77
65.5%; Pred. No. 3.2e-06;
ive 2; Mismatches 8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL, AK083926, BAC39069.1; -.
SEQUENCE 1007 AA, 116975 WW; 18E1B7371E17AB4C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 KH----KHKHKHKGKHKHKHKHKG-----KHKHKHKKK 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         40 KHSKHKKKCHKHSKHK-KHKHSSEEDRDKKHKHKHKHKHKH 77
                              Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     776 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=21000495; PubMed=11090624;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=98054002; PubMed=9393853;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 60.4%;
Best Local Similarity 59.0%;
Matches 23; Conservative
01-MAR-2003 (TrEMBLrel. 23, 01-MAR-2003 (TrEMBLrel. 23, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-JAN-1998 (TrEMBLrel. 05,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19; Conservative
                                                                                                 Pre-mRNA protein kinase.
Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nature 420:563-573(2002)
```

1 КНКНКНКНКСКНКНКНКНКНК 29

ò

RESULT

```
initiator function.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08GUC3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            043487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 14
Q8GUC3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      943487

947043487

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

94704

9470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=AX4;
Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
Gloeckner G., Eichinger C., Parra G., April J.F., Guigo R., Kumpf K.,
Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
"Sequence and Analysis of Chromosome 2 of Dictyostelium.";
Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
EMBL; AL15691; AAL2653.1; -.
BYDOCHETICAL PROCESIN.
SEQUENCE 1461 AA; 169095 MW; A867DA194858EA5C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gapa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=98449945; PubMed=9774672; MEDLINE=98449945; PubMed=9774672; MEDLINE=P8449945; PubMed=9774672; Martinez E., Ge H., Tao Y., Yuan C.X., Palhan V., Roeder R.G.; Martinez E., Ge H., Tao Y., Yuan C.X., Palhan V., Roeder R.G.; Novel cofactors and TFIIA mediate functional core promoter selectivity by the human TAFII150-containing TFIID complex."; Mol. Cell. Biol. 18:657.1:583(1998).
EMBL; AF040701, AAC68502.1; InterPro; IPR001930; Ala peptase.
InterPro; IPR001930; Ala peptase.
Pfam; PF01433; Peptidase M1; 1.
SEQUENCE 1189 AA; 135833 MW; 56EA49A93D9ADA99 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     57.1%; Score 104.5; DB 4; Length 1189; llarity 67.9%; Pred. No. 1.4e-05; Conservative 1; Mismatches 7; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          58 5%; Score 107; DB 5; Length 1461; 41.4%; Pred. No. 7.9e-06;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2; Indels
                                                                                                                                                            01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Hypothetical protein.
Dictyostellum discoideum (Slime mold).
Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.
NCBI_TAXID=44689;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-JUN-1998 (TrEMBLrel. 06, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) TATA binding protein associated factor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 1189 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    41.4%; Pred. ...
         382 KHKHKHKHKHKHNEKRKSKDKDKYKDKYK 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 КНКНКНКНКСКНКНКНКККНККНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 НКНКНКНКСКНКНКНКНКСКНКНКК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-JUN-1998 (TrEMBLrel. 06,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 41.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity
tes 19; Conserv
                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TAFII150.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    043604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 12
060668 .
ID 06066
AC 06066
                                                                                                                                                          Q8ST04
                                                                                                                                   Q8ST04
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Jatches
                                                                                       RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                   셤
```

PRT; 1199 AA.

PRELIMINARY;

060668;

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens (Human).
Bukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDINE=98078679; PubMed=9418870;
Kaufmann J., Ahrens K., Koop R., Smale S.T., Muller R.,
"CIPISO, a human cofactor for transcription factor IID-dependent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 57.1%; Score 104.5; DB 4; Length 1199; 67.9%; Pred. No. 1.4e-05; i.ve 1; Mismatches 7; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                Length 1199;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1199 AA; 136994 MW; 82DA6EAA189D3EF5 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 67.9%; Pred. No. 1.4e-05;
Matches 19; Conservative 1; Mismatches 7; Indels
                                                                                                                                                                                                                                                   Guernah M., Roeder R.G.R.;
Guernah M., Roeder R.G.R.;
"Human TBP-associated factor (TAFII150).";
Submitted (PRR-1998) to the EMBL/GenBank/DDBJ databases.
EMBL; AF057694; AAC13540.1; -.
InterPro; IPR001930; Ala peptase.
Pfam; PF01433; Peptidase M1; 1.
ERQUENCE 1199 AA; 136585 MW; 16F993E6D2378A59 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-JUN-1998 (TrEMBLrel. 06, Created) 01-JUN-1998 (TrEMBLrel. 06, Last sequence update) 01-OCT-2002 (TrEMBLrel. 22, Last annotation update) Cofactor of initiator function.
01-AUG-1998 (TrEMBLrel. 07, Created)
01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
01-AUG-1998 (TrEMBLrel. 22, Last annotation update)
TBP-associated factor TAFII150.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT; 1199 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 НКНКНКНКНКНКНКНКНКНК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2 НҮДІКНІКНІКНІКНІКНІКНІКНІКНІК 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genew; HGNC:11536; TAF2.
InterPro; IPR001930; Ala peptase.
Pfam; PF01433; Peptidase_M1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mol. Cell. Biol. 18:233-239(1998)
EMBL; AF026445; AAC02966.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (TrEMBLrel. 23, C
(TrEMBLrel. 23, I
(TrEMBLrel. 23, I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 57.1
Best Local Similarity 67.9
Matches 19; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hordeum vulgare (Barley)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-MAR-2003 (TrEMBLre
01-MAR-2003 (TrEMBLre
Barley B recombinant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SGETTTRE
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                              SEQUENCE FROM N.A.

STRAIN=cv. Mutant Cal/C15;
Santi L., Wang Y., Mueller K., Stile M.R., Roig C., Wanke D.,
Santi L., Mang Y., Mueller J., Pozzi C., Rohde W., Salamini F.;
"Cloning a plant (GC/TC)8 binding factor supports an intron mediated
model for plant homeobox gene regulation.";
Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.

EMBL; AJ507214; CAD45251.1;
SEQUENCE 350 AA; 38978 MW; 668DF5294A12E061 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0; Gaps
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Lillopsida; Poales; Poaceae; Pooideae;
Triticeae; Hordeum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F., Prensier G., Barbe V., Peyretaillade E., Brottier P., Wincker P., Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M., Weissenbach J., Vivares C.P.;
"Genome sequence and gene compaction of the eukaryote parasite Encephalitozoon cuniculi.";

Mature 414:50-453(2001).

EMBL; ALS90448; CAD26474.1; -.

Hypotherical protein.

SEQUENCE 260 AA; 29635 MW; A307DD4D452F4E82 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
Hypochetical protein ECU08_1700.
ECU08_1700.
Encephalitozoon cuniculi.
Eukaryota; Pungi; Microsporidia; Unikaryonidae; Encephalitozoon.
                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                       Query Match 56.8%; Score 104; DB 10; Length 350; Best Local Similarity 42.9%; Pred. No. 4.8e-06; Matches 12; Conservative 14; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    54.6%; Score 100; DB 5; Length 260; 53.6%; Pred. No. 1.2e-05; tive 3; Mismatches 10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            260 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 КНКНКНКНКСКНКНКНКНКСКНКНКН 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 КНКНКНКНКСКНКНКНКНКНКНКНКНКН 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN=GB-M1;
MEDLINE=21576510; PubMed=11719806;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 53.61
Matches 15; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. STRAIN=GB-M1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                  NCBI_TaxID=4513;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genoscope;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q8SUK4
Q8SUK4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 15
Q8SUK4
    SOCKERFFFF
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠે
```

a

Search completed: January 20, 2004, 18:26:21 Job time : 37.7246 secs

	7	7/	A	3	A S							
											•	
						S- 12 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1						
						•						
				- 1,	· ·		·				 •	
						•						
	* -											
			1.			**						
	e * ' .											·
							. **					
0.												
						. *	·					
												e8. Y
										•		
-												• ,
				*								*
						٠						
							•					÷
		1										
		•										
						•						
							•					
٠												
				•								
											. =	
				. •	, .				 ٠			
		i,								:		
			الا. غار		T.J		* **		3			
			790		4 36 1			Tp :				
			4			See. 17.						-0.0
			e e	7	4			-10 3				14 h
					*			•				The state of the s

```
January 20, 2004, 18:05:34; Search time 31.0145 Seconds (without alignments) 102.356 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A censer 250 Millo 1.

1. (SIDSI) gradata/geneseq/geneseqp-embl/AA1980.DAT:*

2. (SIDSI) gradata/geneseq/geneseqp-embl/AA1981.DAT:*

3. (SIDSI) gradata/geneseq/geneseqp-embl/AA1991.DAT:*

4. (SIDSI) gradata/geneseq/geneseqp-embl/AA1991.DAT:*

5. (SIDSI) gradata/geneseq/geneseqp-embl/AA1991.DAT:*

5. (SIDSI) gradata/geneseq/geneseqp-embl/AA1991.DAT:*

6. (SIDSI) gradata/geneseq/geneseqp-embl/AA1991.DAT:*

7. (SIDSI) gradata/geneseq/geneseqp-embl/AA1991.DAT:*

8. (SIDSI) gradata/geneseq/geneseqp-embl/AA1992.DAT:*

7. (SIDSI) gradata/geneseq/geneseqp-embl/AA1992.DAT:*

8. (SIDSI) gradata/geneseq/geneseqp-embl/AA1992.DAT:*

8. (SIDSI) gradata/geneseq/geneseqp-embl/AA1992.DAT:*

8. (SIDSI) gradata/geneseq/geneseqp-embl/AA1999.DAT:*

8. (SIDSI) gradata/geneseq/geneseqp-embl/AA1999.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1107863
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1107863 segs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                               OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                136
1 КНКНИКНКНКНККНК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A Geneseq 19Jun03:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                  US-10-018-103A-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                            Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Database
                                                                                                                                                                                                                   Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	Histidine copolyme	HH-K peptide segue	HH-K2b polymer. U	HH-K3b polymer. U	HH-K4b polymer. U	Histidine copolyme	Histidine copolyme	Drosophila melanog	Novel human diagno
SUMMARIES	QI	AAE06230	AAE06249	AAE06245	AAE06246	AAE06247	AAE06233	AAE06240	ABB68923	ABG14399
	DB	22	22	22	22	22	22	22	22	22
	Match Length DB	20	20	41	62	83	26	15	290	227
* Ouery	Match	100.0	100.0	100.0	100.0	100.0	80.1	77.2	71.3	69.1
	Score	136	136	136	136	136	109	105	97	94
Result	No.	-	7	m	4	S	9	7	80	σ

phila r like #465 in #458 in #	Novel human diagno Novel human diagno Plant SDF encoded Arabidopsis thalia Novel human diagno Novel human diagno
ABB60857 ABB60854 ABG52113 ABB32046 ABB37295 AAM70433 AAM70433 AAM70433 AAM6572 AAM6572 AAM6572 AAM6572 AAM6572 AAM6572 AAM6572 AAM70433 AAM6572 AAM99836 AAM99836 AAM99836 AAM99836 AAM99838	ABG05192 ABG25337 AAB24811 AAG44771 ABG28019 ABG03494
11,000,000,000,000,000,000,000,000,000,	22 22 23 23 24 25 25 27 27 27 27 27 27 27 27 27 27 27 27 27
	134 193 195 245 266
000000000000000000000000000000000000000	
4 4 W W U U U U U W W W W W W W W W W W	22222 20222
01111111110222222222222222222222222222	4 4 4 4 4 5 5 4 6 7 1 6

ALIGNMENTS

familial hypercholesterolaemia; low-density lipoprotein receptor; phenylketonuria; Faconi's anaemia; haemophilia; muscular dystrophy; cystic fibrosis; Parkinson's disease; retinitis pigenthosa; cycostatic; lysosomal storage disease; mycopolysaccharide type I disease; cardiant; diabetic retinopathy; human immunodeficiency virus disease; infection; cardiac disease; peripheral vascular disease; antisickling; arthritis; Histidine copolymer; enzyme deficiency; phenylalanine hydroxylase; ornithine transcarbamylase; adenainase; tyroshine hydroxylase; purine mucleoside phosphorylase; alphal-antitrypsin; apolipoprotein E; chronic granulomatous disease; sickle cell; beta-thalassaemia; anaemia; nootropic; haemostatic; virucide; gene therapy Histidine copolymer, HH-K peptide. AAE06230 standard; peptide; 20 AA. 25-SEP-2001 (first entry) AAE06230; AAE06230

Unidentified

WO200147496-A1

20-DEC-2000; 2000WO-US34603. 05-JUL-2001

99US-0173576 29-DEC-1999;

(MIXS/) MIXSON A J

```
AAE06249 standard; peptide; 20 AA
                                                                                                                                                            1 KHKHHKHKHKHKHKKK 20
                                                                                                                                                        1 КНКНИКНИКНИКНКИК 20
                                     Claim 3; Page 36; 64pp; English.
                                                                                                                                             Local Similarity 100.
nes 20; Conservative
                                                                                                                                                                                                      HH-K peptide sequence.
             WPI; 2001-425579/45
                                                                                                                                                                                               25-SEP-2001
                                                                                                                                                                                                                                                      Unidentified
                                                                                                                                    Sequence
                                                                                                                                                                                        AAE06249;
                                                                                                                                           Query Match
                                                                                                                                                  Matches
                                                                                                                                                                          RESULT 2
                                                                                                                                                                              AAE06249
                                                                                                                                                                                     8
                                                                                                                                                               g
```

20 AA;

100.0%; Score 136; DB 22; 100.0%; Pred. No. 1.1e-11; ative 0; Mismatches 0;

ö

(first entry)

```
The invention relates to a pharmaceutical agent delivery composition comprising a transport polymer comprising a linear or branched peptide comprising a transport polymer comprising a linear or branched peptide having about 10 to 300 amino acid residues, having about 5 to 100% histidine residues, and optionally having 10 95% non-histidine listed to be a least one pharmaceutical agent and optionally one or more intracellular delivery composition is useful for cellivering a pharmaceutical agent to the interior of a cell and is be a pharmaceutical agent to the interior of a cell and is cellivering a pharmaceutical agent to the interior of a cell and is considered bhosphorylase deficiency, chronic granulomatous disease with nucleoside phosphorylase deficiency, chronic granulomatous disease with condequate production of beta-hamengolobin, familial hypercholesterolaemia due to a defective payphox, sackle cell with HbS, beta-thalassaemia due to confithine transcarbamylase deficiency, raconi's anaemia, apolipoprocein confithine transcarbamylase deficiency, raconi's anaemia, apolipoprocein confithine transcarbamylase deficiency, raconi's anaemia, apolipoprocein confithine dystrophial and B due to factor VIII and IX deficiency, muscular dystrophin, laminin-2, or sacroalycans cutations, cystic fibrosis due to CFTR mutations, parkingon due to tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage cyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage confidencial cretinopathy, human immundeficiency virus disease, virus in HH-K2b, HH-K3b and HHK4b polymers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Enzyme deficiency; phenylalanine hydroxylase; ornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; purine nucleoside phosphorylase; alphal-antitrypsin; apolipoprotein E; chronic granulomatous disease; sickle cell; beta-thalassasemia; anaemia; familial hypercholesteroleamia; low-density lipoprotein receptor; phenylketonuria; Faconi's anaemia; haemophilia; muscular dystrophy;
                                                                                                                                                                                                                                                                                                                                                                                                                           pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100.0%; Score 136; DB 22;
100.0%; Pred. No. 1.1e-11;
iive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAE06245 standard; peptide; 41 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Disclosure; Fig 17; 64pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 KHKHHKHHKHHKHKHK 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 KHKHKHKHKHKHKHK 20
                                                                                                                  20-DEC-2000; 2000WO-US34603
                                                                                                                                                                                     99US-0173576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity 100.
nes 20; Conservative
                                                                                                                                                                                                                                                                                                                                                                           WPI; 2001-425579/45.
                                                                                                                                                                                                                                                  (MIXS/) MIXSON A J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HH-K2b polymer.
WO200147496-A1
                                                                                                                                                                                         29-DEC-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25-SEP-2001
                                                            05-JUL-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAE06245;
                                                                                                                                                                                                                                                                                                                     Mixson AJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAE06245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                          The invention relates to a pharmaceutical agent delivery composition comprising a transport polymer comprising a linear or branched peptide having about 10 to 300 amino acid residues, having about 5 to 100% having about 10 to 300 amino acid residues, having about 5 to 100% having about 10 to 300 amino acid residues, having 1 to 95% non-histidine residues, at least one pharmaceutical agent and optionally one or more intracellular delivery components in association with the transport colymer. The pharmaceutical agent delivery composition is useful for delivering a pharmaceutical agent to the interior of a cell and is used in gene therapy for treating adenosine deaminase deficiency, purine cuclosside phosphorylase deficiency, chronic granulomatous disease with necessity and the HBS, beta-thalassaemia due to inadequate production of beta-hemoglobin, familial hypercholesterolaemia cor a defective low-density lipoprotein receptor, phenylkecountia due to a defective low-density lipoprotein receptor, phenylkecountia due to phenylalanine hydroxylase deficiency, Paconi's anaemia, apolipoprotein cor phenylalania dystrophida de to factor VIII and IX deficiency, muscular dystrophida due to dystrophin, laminin-2, or sacroglycans cuscular dystrophida due to GTR mutations, cystic fibrosis due to GTR mutations, Parkinson due to tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage confisease (i.e., mycopolysaccharide type I, Hunter, Hurler and Gaucher), diabetic retinopathy, human immunodeficiency virus disease, virus infection, acquired anaemia, cardiac and peripheral vascular disease infection, acquired anaemia, cardiac and peripheral vascular disease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Enzyme deficiency; phenylalanine hydroxylase; cornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; purine nucleoside phosphorylase; alphal-antitrypsin; apolipoprotein E; chronic granulomatous disease; sickle cell; beta-thalassaemia; anaemia; familial hypercholesterolaemia; low-density lipoprotein receptor; phenylketohuria; Faconi's anaemia; haemophilia; muscular dystrophy; cystic fibrosis; parkinson's disease; retinitis pigmentosa; cytostatic; lysosomal storage disease; mycopolysaccharide type I disease; cardiant; diabetic retinopathy; human immundeficiency virus disease; infection; cardiac disease; peripheral vascular disease; antisickling; arthritis; nootropic; haemostatic; virucide; gene therapy.
                                                                                                                                                                                            Pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0; Indels
```

ö

Gaps ; 0

```
The invention relates to a pharmaceutical agent delivery composition comprising a transport polymer comprising a linear or branched peptide having about 1 to to 300 amino acid residues, having about 5 to 100% thistidine residues, and optionally having 1 to 95% non-histidine residues, and optionally having 1 to 95% non-histidine intracellular delivery components in association with the transport polymer. The pharmaceutical agent delivery composition is useful for delivering a pharmaceutical agent delivery composition is useful for used in gene therapy for treating adenosine deaminase deficiency, purine uncleoside phosphorylase deficiency, chronic granulomatous disease with defective p47phox, sickle cell with HbS, beta-thalassaemia due to inadequate production of beta-haemoglobin, familial hypercholesterolaemia contithine transcarbamylase deficiency, alphal-antitrypsin deficiency, ornithine transcarbamylase deficiency, alphal-antitrypsin deficiency, contithine transcarbamylase deficiency, alphal-antitrypsin deficiency, muscular dystrophy due to dystrophin, laminn-2, or sacroglycans chieses (i.e., mycopolysaccharide type 1, Hunter, Hurlar and Gaucher), diabetic retinopathy, human immunodeficiency virus disease, virus infection, acquired anaemia, cardiac and peripheral vascular disease infection, acquired anaemia, cardiac and peripheral vascular disease
cystic fibrosis; Parkinson's disease; retinitis pigmentosa; cytostatic;
                      lysosomal storage disease, mycopolysaccharide type I disease; cardiant; diabetic retinopathy; human immunodeficiency virus disease; infection; cardiac disease; peripheral vascular disease; antisickling; arthritis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues
                                                                                             nootropic; haemostatic; virucide; gene therapy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example; Page 23; 64pp; English.
                                                                                                                                                                                                                                                                                     20-DEC-2000; 2000WO-US34603
                                                                                                                                                                                                                                                                                                                                 99US-0173576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2001-425579/45.
                                                                                                                                                                                                                                                                                                                                                                             (MIXS/) MIXSON A J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       41 AA;
                                                                                                                                                                                     WO200147496-A1.
                                                                                                                                                                                                                                                                                                                               29-DEC-1999;
                                                                                                                                         Unidentified
                                                                                                                                                                                                                                    05-JUL-2001
                                                                                                                                                                                                                                                                                                                                                                                                                               Mixson AJ;
```

ö Gaps ö Length 41; Indels 100.0%; Score 136; DB 22; 100.0%; Pred. No. 2.3e-11; tive 0; Mismatches 0; Query Match
Best Local Similarity 100.0
Matches 20; Conservative

ઠે

RESULT 4

25-SEP-2001 (first entry) HH-K3b polymer

AAE06246 standard; peptide; 62 AA AAE06246; *****

Enzyme deficiency; phenylalanine hydroxylase; cornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; ornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; purine nucleoside phosphorylase; alphal-antitrypsin; apolipoprotein E; chronic granulomatous disease; sickle cell; beta-thalassaemia; anaemia; familial hypercholesterolaemia; low-denosity lipoprotein receptor; phenylketomuria; Paconi's anaemia; haemophila; muscular dystrophy; cystic fibrosis; Parkinson's disease; retinitis pigmentosa; cytostatic; lysosomal storage disease; mycopolysaccharide type I disease; cardiant; diabetic retinopathy; human immunodeficiency virus disease; infection; cardiac disease; peripheral vascular disease; antisickling; arthritis; Pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues nootropic; haemostatic; virucide; gene therapy. 20-DEC-2000; 2000WO-US34603. 99US-0173576 WPI; 2001-425579/45. (MIXS/) MIXSON A J. WO200147496-A1. 29-DEC-1999; Unidentified. 05-JUL-2001. Mixson AJ;

The invention relates to a pharmaceutical agent delivery composition comprising a transport polymer comprising a linear or branched peptide comprising about 10 to 300 andno acid residues, having about 5 to 100% thistidine residues, and optionally having 1 to 95% non-histidine residues, at least one pharmaceutical agent and optionally one or more intracellular delivery components in association with the transport components in association with the transport polymer. The pharmaceutical agent delivery composition is useful for delivering a pharmaceutical agent delivery composition is useful for used in gene therapy for treating adenosine deaminase deficiency, purine cused in gene therapy for treating adenosine deaminase deficiency, purine to the defective p47phox, sickle cell with HbS, beta-thalassaemia due to defective p47phox, sickle cell with HbS, beta-thalassaemia due to inadequate production of beta-haemoglobin, familial hypercholesterolaemia come to a defective low-density lipoprotein receptor, phenylalanine hydroxylase deficiency, alphal-antitrypsin deficiency, ornithine transcarbamylase deficiency, Faconi's anaemia, apolipoprotein E deficiency, haemophilia A and B due to factor VIII and IX deficiency, muscular dystrophy due to dystrophin, laminin-2, or sacrogly-cans cuscular dystrophy when immunodeficiency virus disease, virus clabetic retinopathy, human immunodeficiency virus disease, virus infection, acquired anaemia, cardiac and peripheral vascular disease condition, acquired anaemia, cardiac and peripheral vascular disease.

Example, Page 23; 64pp; English.

62 AA; Sequence

Gaps ; Length 62; Indela 22; Score 136; DB 22; Pred. No. 3.6e-11; ; Mismatches 0; ö 100.08; 100.001 Query Match Best Local Similarity 100. Matches 20; Conservative

ö

Š 셤 RESULT 5 AAE06247

```
AAE06247 standard; peptide; 83 AA
```

AAE06247;

25-SEP-2001 (first entry)

HH-K4b polymer.

Enzyme deficiency; phenylalanine hydroxylase; ornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; purine nucleoside phosphorylase; alphal-antitrypsin; apolipoprotein E; chronic granulomatous disease; sickle cell; beta-thalassaemia; anaemia; familial hypercholesterolaemia; low-density lipoprotein receptor; phenylketonuria; Faconi's anaemia; heemophilia; muscular dystrophy; cystic fibrosis; Parkinson's disease; retinitis pigmentosa; cytostatic; lysosomal storage disease; mycopolysaccharide type I disease; cardiant; diabetic retinopathy; human immundeficiency virus disease; infection; cardiac disease; peripheral vascular disease; antistickling; arthritis; nootropic; haemostatic; virucide; gene therapy.

Jnidentified

WO200147496-A1.

05-JUL-2001.

20-DEC-2000; 2000WO-US34603.

99US-0173576. 29-DEC-1999;

(MIXS/) MIXSON A J.

Mixson AJ;

WPI; 2001-425579/45.

Pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues

Example; Page 23; 64pp; English.

The invention relates to a pharmaceutical agent delivery composition comprising a transport polymer comprising a linear or branched peptide comprising a transport polymer comprising a linear or branched peptide having about 10 to 300 amino acid residues, having about 5 to 100% institution residues, and optionally having 1 to 95% non-histidine measured to the pharmaceutical agent and optionally one or more contracellular delivery components in association with the transport of polymer. The pharmaceutical agent delivery composition is useful for delivering a pharmaceutical agent delivery composition is useful for used in gene therapy for treating adenosine deaminase deficiency, purine conclosate production of beta-hemoglobin, familial hypercholesterolaemia continue production of beta-hemoglobin, familial hypercholesterolaemia due to a defective low-density lipoprotein reseptor, phenylkeronuria due continine transcarbamylase deficiency, alphal-antitrypsin deficiency, continine transcarbamylase deficiency, panhal-antitrypsin deficiency, muscular dystrophilia A and B due to factor VIII and IX deficiency, muscular dystrophilia A and B due to factor VIII and IX deficiency, contininal dystrophilia A and B due to factor VIII and IX deficiency, muscular dystrophylase deficiency, retinitis pigmentosa, lysosomal storage cyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage cyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage cyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage confisease (i.e., mycopolysaccharide type 1, Hunter, Hurler, Hurler, diaberic retinopathy, human immunodeficiency virus disease, virus infection, acquired anaemia, cardiac and peripheral vascular disease infection, acquired enaemia, cardiac and peripheral vascular disease.

Sequence

```
ö
                                               Gaps
                                                ö
                    Length 83;
                                                Indels
                     100.0%; Score 136; DB 22;
100.0%; Pred. No. 4.8e-11;
Live 0; Mismatches 0;
                           Query Match
Best Local Similarity 100.
Matches 20; Conservative
83 AA;
```

```
Histidine copolymer; enzyme deficiency; phenylalanine hydroxylase; ornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; purine nucleoside phosphorylase; alphal-antitrypsin; apolipoprotein E; familial granulomatous disease; sickle cell; beta-thalassaemia; anaemia; familial hypercholesterolaemia; low-density lipoprotein receptor; phenylketconuria; Faconi's anaemia; haemophilia; muscular dystrophy; cystic fibrosis; Parkinson's disease; relinitis pigmentosa; cytostatic; lysosomal storage disease; mycopolysaccharide type I disease; cardiant; diabetic retinopathy; human immunodeficiency virus disease; infection; cardiac disease; peripheral vacular disease; antisickling; arthritis;
                                                                                                                                                                                                                                                 nootropic; haemostatic; virucide; gene therapy.
                                                                                                                         Histidine copolymer, Y-HH peptide.
                                                            AAE06233 standard; peptide; 26 AA
1 KHKHHKHKHKHKHKHK 20
                                                                                                      25-SEP-2001 (first entry)
                                                                                                                                                                                                                                                                         Unidentified
                                                                                   AAE06233;
                                                   AAE06233
                                                                          g
```

/note= "Y component" WO200147496-A1. 05-JUL-2001. Key Region

Location/Qualifiers

20-DEC-2000; 2000WO-US34603

99US-0173576. 29-DEC-1999;

(MIXS/) MIXSON A J.

Mixson AJ;

WPI; 2001-425579/45.

Pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues

Example; Page 23; 64pp; English.

The invention relates to a pharmaceutical agent delivery composition comprising a transport polymer comprising a linear or branched peptide comprising a transport polymer comprising a linear or branched peptide comprising about 5 to 100% and optionally having 1 to 95% non-histidine residues, and optionally having 1 to 95% non-histidine calcius, at least one pharmaceutical agent and optionally one or more residues, at least one pharmaceutical agent delivery composition with the transport composition with the transport conjumer. The pharmaceutical agent delivery composition is useful for delivering a pharmaceutical agent delivery composition is useful for concleoside phosphorylase deficiency, chronic granulomatous disease with confective p47phox, sickle cell with HbS, beta-thalasseemia due to defective p47phox, sickle cell with HbS, beta-thalasseemia due to confective p47phox, sickle cell with HbS, beta-thalasseemia due to confective p47phox, sickle cell with HbS, beta-thalasseemia due to confective p47phox, alckle cell with HbS, beta-thalasseemia due to confective p47phox, alckle cell with HbS, beta-thalasseemia due to confective p47phox, alckle cell with HbS, beta-thalasseemia due to confective p47phox, alckle cell with HbS, beta-thalasseemia due to confective p47phox, alckle cell with HbS, beta-thalasseemia due to conferince confective p47phox, alckle celloncy, Faconi's anaemia, apolipoprotein confectioncy, hemosphila A and B due to factor VIII and IX deficiency, muscular dystrophy due to dystrophin, laminin-2, or sacroglycans containe hydroxylase deficiency, retinitis pigmentosa, lysosomal storage tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The invention relates to a pharmaceutical agent delivery composition comprising a transport polymer comprising a linear or branched peptide having about 10 to 300 amino acid residues, having about 5 to 100% histidine residues, and optionally having 1 to 95% non-histidine residues, at least one pharmaceutical agent and optionally one or more intracellular delivery components in association with the transport polymer. The pharmaceutical agent to the interior of a cell and is delivering a pharmaceutical agent to the interior of a cell and is nucleoside phosphorylase deficiency, chronic granulomatous disease with defective p47phox, sickle cell with HbS, beta-thalassaemia due to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Histidine copolymer; enzyme deficiency; phenylalanine hydroxylase; ornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; purine mucleoside phosphorylase; alphal-antitrypsin; apolioporotain B; chronic granulomatous disease; sickle cell; beta-thalassaemia; anaemia; familial hypercholesterolaemia; low-density lipoprotein receptor; phenylkeronuria; Paconis anaemia; heemophilia; muscular dystrophy; cystic fibrosis; Parkinson's disease; retinitis pigmentosa; cytostatic; lysosomal storage disease; mycopolysaccharide type I disease; cardiant; diabetic retinopathy; human immunodeficiency virus disease; infection; cardiac disease; peripheral vascular disease; antisickling; arthritis; nootropic; haemostatic; virucide; gene therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues
 mycopolysaccharide type 1, Hunter, Hurler and Gaucher),
                   diabetic retinopathy, human immunodeficiency virus disease, virus infection, acquired anaemia, cardiac and peripheral vascular disease and arthritis. The present peptide sequence is histidine copolymer
                                                                                                                                                                                          Gaps
                                                                                                                                                                                         ô
                                                                                                                                                       Length 26;
                                                                                                                                                                                         Indels
                                                                                                                                                 Match 80.1%; Score 109; DB 22; Local Similarity 83.3%; Pred. No. 4.5e-08; les 15; Conservative 2; Mismatches 1;
                                                                                                                                                                                                                                                                                                                                                                  Ą.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 3; Page 36; 64pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Histidine copolymer, peptide #3.
                                                                                                                                                                                                                                                                                                                                                                AAE06240 standard; peptide; 15
                                                                                                                                                                                                                              1 КИКНИКНИКНИКНИКНИК 18
                                                                                                                                                                                                                                                                26
                                                                                                                                                                                                                                                  20-DEC-2000; 2000WO-US34603.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0173576.
                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2001-425579/45.
                                                                                                               26 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200147496-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NOSXIW (/SXIW)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29-DEC-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                             25-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             05-JUL-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mixson AJ;
                                                                                                                                                                                                                                                                                                                                                                                                   AAE06240;
                                                                                                                 Sequence
                                                                                                                                                     Query Match
                                                                                                                                                                                         Matches
                                                                           Y-HH.
                                                                                                                                                                                                                                                                                                                            RESULT 7
                                                                                                                                                                                                                                                                                                                                                AAE06240
88888888
                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                     ઠ
```

```
ö
inadequate production of beta-haemoglobin, familial hypercholesterolaemia
             due to a defective low-density lipoprotein receptor, phenylketonuria due to phenylalanine hydroxylase deficiency, alphal-antitrypsin deficiency, ornithine transcarbamylase deficiency, Faconi's anaemia, apolipoprotein E deficiency, haemophilia A and B due to factor VIII and IX deficiency, muscular dystrophy due to dystrophin, laminin-2, or sacrodycans mutations, cystic fibrosis due to CFTR mutations, Parkinson due to tyrosine hydroxylase deficiency, retinitis pigmentosa, lysosomal storage disease (i.e., mycopolysaccharide type 1, Hunter, Hurler and Gaucher), diabetic retinopathy, human immunodeficiency virus disease, virus infection, acquired anaemia, cardiac and peripheral vascular disease and arthritis. The present peptide sequence is a histidine copolymer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01875) and the encoded proteins
                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; SEQ ID NO 33561; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                          Length 15;
                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Drosophila melanogaster polypeptide SEQ ID NO 33561.
                                                                                                                                                                                                                                                                        Score 105; DB 22;
Pred. No. 8.4e-08;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Myers EW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABB68923 standard; Protein; 290 AA.
                                                                                                                                                                                                                                                                                        100.0%; Pre
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Li PWD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23-MAR-2001; 2001WO-US09231.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
                                                                                                                                                                                                                                                                                                                                                   4 НИКНИКИМИТИТИТИ 18
                                                                                                                                                                                                                                                                                                                                                                                    1 никникникник 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                 Query Match ''...
Best Local Similarity 100.'
Matches 15; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Drosophila melanogaster.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Venter JC, Adams M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2001-656860/75.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (PEKE ) PE CORP NY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2001-656860/
N-PSDB; ABL13026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200171042-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     pharmaceutical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             27-SEP-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Drosophila;
                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABB68923;
                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 8
ABB68923
88888888888888888888
                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠ
```

Length 227;

g

셤 ò

```
Venter JC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Seguence
                                                                                                                                                                   ABB60857;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 11
                                                                                                               RESULT 10
                                                                                                                            ABB60857
                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                   셤
                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The invention relates to isolated polymuclectide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PGR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The and gene mapping, and in recombinant production of (II). The polymucleotides are also used in diagnostics as expressed sequence tags polymucleotides acrivity of (II) or to treat disease states involving to respect the polymer of (II) and its useful for generating antibodies against it, detecting or (II). (II) is useful for generating antibodies against it, detecting or supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating imaging of sites expressing (II). (I) and (II) are useful for treating clisorders involving aberrant protein espression or biological activity. The proposible for genetic disorders or other traits to assess biodiversity responsible for genetic disorders or other traits to assess biodiversity amino acid sequence other types of data and products dependent on DNA and amino acid sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from MIPO at fire, wipo.int/pub/published_pct_sequences.
                                              ;
0
                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New isolated polynuclectide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity
                                                                                                                                                                                                                                                                                 Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
                                              ö
                    Length 290;
                                                 Indels
                  Score 97; DB 22; L4
Pred. No. 1.8e-05;
1; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 20; SEQ ID No 44758; 103pp; English.
                                                                                                                                                                                                                                                          Novel human diagnostic protein #14390
                                                                                                                                                                           ABG14399 standard; Protein; 227 AA.
                                                                                            || || || || || || || || 118 || 136
                                                                            1 КНКНИКНИКНИМ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Tang YT;
                                                                                                                                                                                                                                                                                                                                                                                                                                        31-MAR-2000; 2000US-0540217.
23-AUG-2000; 2000US-0649167.
                                                                                                                                                                                                                                                                                                                                                                                                              30-MAR-2001; 2001WO-US08631
                       Query Match
71.3%;
Best Local Similarity 68.4%;
Matches 13; Conservative
                                                                                                                                                                                                                                18-FEB-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Drmanac RT, Liu C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2001-639362/73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  227 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (HYSE-) HYSEQ INC
290 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAS78586
                                                                                                                                                                                                                                                                                                                                                          MO200175067-A2.
                                                                                                                                                                                                                                                                                                                                Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                      11-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                       ABG14399;
Sequence
                                                                                                                                                   RESULT 9
                                                                                                                                                                 ABG14399
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL16175) and the encoded proteins
                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Drosophila; developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5; Indels 0;
                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 94; DB 22; Length 1413;
Pred. No. 0.00022;
2; Mismatches 5; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ID NO 9363; 21pp + Sequence Listing; English.
                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                      Drosophila melanogaster polypeptide SEQ ID NO 9363
Score 94; DB 22;
Pred. No. 3.5e-05;
2; Mismatches 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Li PWD, Myers EW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ABB60854
ID ABB60854 standard; Protein; 1424 AA.
                                                                                                                                                                                                                                                                           ABB60857 standard; Protein; 1413 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1321 OHOHHMHSHHMHPH 1339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 КНКНИКНИКНИМ 19
                                                 2;
                                                                                                                              141 HHHHHHHHHQHHHHQH 158
                                                                                                   2 НИНКНИКНИКНИМ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 69.1%;
Best Local Similarity 63.2%;
Matches 12; Conservative ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     23-MAR-2001; 2001WO-US09231.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
                                                                                                                                                                                                                                                                                                                                                                               26-MAR-2002 (first entry)
     Query Match
Best Local Similarity 66.7
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Drosophila melanogaster.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Adams M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2001-656860/75.
N-PSDB; ABL04960.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (ABB57737-ABB72072)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1413 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (PEKE ) PE CORP NY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; SEQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             pharmaceutical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200171042-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              27-SEP-2001.
```

```
The invention relates to a single exon nucleic acid probe (SENP) (I) for measuring human gene expression in a sample derived from human adult
liver, comprising one of 13109 defined nucleotide sequences given in the
specification (or complements/ fragments). The probe hybridises at high
stringency to a nucleic acid molecule expressed in the human adult
liver. (I) may be used for predicting, measuring and displaying gene
cxpression in samples derived from human adult liver. The genes
cidentified may be involved in genetic liver diseases such as cirrhosis,
hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
is associated with coronary heart disease. ABG47348-ABG59930 represent
human liver single exon encoded peptides of the invention.
Note: The sequence information for this patent does not appear in the
printed specification but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Peptide #4697 encoded by breast cell single exon nucleic acid probe.
                                                                                                                                                                                                                                                                                                                                                 Human genome-derived single exon nucleic acid probes useful for analysing gene expression in human adult liver -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Human; microarray; single exon probe; gene expression; breast;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 93; DB 22; Length 49;
Pred. No. 9.8e-06;
1; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                         Claim 27; SEQ ID No 30761; 658pp; English
                                                                                                                                                                                                                                                                        Chen W, Rank DR,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ABB32046 standard; Peptide; 49 AA.
                                                                                                                                                                                                                                  (MOLE-) MOLECULAR DYNAMICS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 НКНИКНИКНИКН 19
                                                                                                                                    2000US-0632366.
2000US-0234687.
2000US-0236359.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  68.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  04-PEB-2000; 2000US-0180312.
26-MAY-2000; 2000US-0207456.
30-JUN-2000; 2000US-0608408.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 нининининининин
                                      30-JAN-2001; 2001WO-US00664
                                                                                              2000US-0207456
2000US-0608408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30-JAN-2001; 2001WO-US00662
                                                                                                                                                                                            04-OCT-2000; 2000GB-0024263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 66.74
Sept. Local 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                        Penn SG, Hanzel DK,
                                                                                                                                                                                                                                                                                                               WPI; 2001-488898/53.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              49 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           disease; cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO200157271-A2
                                                                                                                                    03-AUG-2000; 2
21-SEP-2000; 2
27-SEP-2000; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-FEB-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens
                                                                                                                    30-JUN-2000;
                                                                                              26-MAY-2000;
  09-AUG-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ABB32046;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ABB3204
  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         solated nucleic acid detection reagent for detecting 1000 or more from Drosophila and for elucidating cell signalling and cell-cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention useful in developmental biology and in elucidating cell signalling and insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL16176-ABL30511), expressed DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia; coronary heart disease.
                                                                                                                developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; SEQ ID NO 9354; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   69.1%; Score 94; DB 22; Length 1424; 63.2%; Pred. No. 0.00023; ive 2; Mismatches 5; Indels (
                                                                              Drosophila melanogaster polypeptide SEQ ID NO 9354
                                                                                                                                                                                                                                                                                                                                                                                                                               Myers EW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human liver peptide, SEQ ID No 30761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ABG52113 standard; Peptide; 49 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1332 QHQHRMHSHHMHAHPH 1350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 КНКНИКНИКНИКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                               Li PWD,
                                                                                                                                                                                                                                                                                          23-MAR-2001, 2001WO-US09231.
                                                                                                                                                                                                                                                                                                                               23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity 63.2
Les 12; Conservative
                                                                                                                                                                          Drosophila melanogaster.
                                                                                                                                                                                                                                                                                                                                                                                                                               Adams M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2001-656860/75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (ABB57737-ABB72072)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1424 AA;
                                                                                                                                                                                                                                                                                                                                                                                     (PEKE ) PE CORP NY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-PSDB; ABL04957
                                                                                                                                                                                                              WO200171042-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 interactions -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO200157273-A2
                                                                                                                                      pharmaceutical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens
                                      26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                               Venter JC,
                                                                                                                  Drosophila;
                                                                                                                                                                                                                                                      27-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Si
Matches 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
ABB60854;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ABG52113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genes
```

RESULT 12

a

ઠે

ö

Gaps

. 0

```
Penn SG,
                                                                                                                                                          Human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABB22584
ID ABB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ଟ
                                                                                                                                                                                                                                                                                               The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived nucleic acid probes for measuring gene expression in a sample derived from human breast and BT 474 cells. The method involves contacting the probes with a collection of detectably labelled nucleic acids derived from mRNA of human breast, and then measuring the label bound to each probe of the microarray. The probes are useful for verifying the expression of regions of genomic DNA predicted to encode proteins. They are useful for gene discovery, and for determining predisposition and/or prognosing breast disease. Gene expression analysis is useful for assessing the toxicity of chemical agents on cells. The microarray of this invention presents a far greater diversity of probes for measuring gene expression, with far less bias diversity of probes for measuring gene expression, with far less bias than expressed sequence tay microarrays. The method is sultable for rapid production of functional information from genomic sequence. The probes are presents expensed acid as peptide encoded by a single exon nucleic acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           probe of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Human; foetal liver; gene expression; single exon nucleic acid probe.
                                                                                                                                                                                              New spatially-addressable set of single exon nucleic acid probes, useful for measuring gene expression in sample derived from human breast, comprises number of single exon nucleic acid probes -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Peptide #4801 encoded by human foetal liver single exon probe.
                                                                                                                                                                                                                                                                    Claim 27; SEQ ID NO 15014; 327pp + sequence listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 93; DB 22; Length 49;
Pred. No. 9.8e-06;
1; Mismatches 5; Indels
                                                                                                                                Rank DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ABB37295 standard; Peptide; 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 нкинкинкинки 19
                                                                                               (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                  Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    68.4%;
66.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      04-FEB-2000; 2000US-0180312.
26-MAY-2000; 2000US-020456.
30-UIN-2000; 2000US-0608408.
03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-0234687.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        30-JAN-2001; 2001WO-US00669
       03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0236359.
04-OCT-2000; 2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         04-FEB-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               es.4
Local Similarity 66.7
Les 12, Conservative
                                                                                                                                    ΩĶ,
                                                                                                                                                                    WPI; 2001-496933/54
                                                                                                                                    Hanzel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200157277-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABB37295;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                    Penn SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ABB37295
```

ò 셤

```
Protein #4583 encoded by probe for measuring heart cell gene expression.
                                                                                                                                                                                                                                                                                                                                               The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human fetal liver. The present sequence is a peptide encoded by a single exon nucleic acid probe of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Single exon nucleic acid probes for analyzing gene expression in human hearts -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gape
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                       genome-derived single exon nucleic acid probes useful for
                                                                                                                                                                                                                                                                                                    Claim 27; SEQ ID NO 29930; 639pp + sequence listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Human, gene expression, heart, microarray, vascular system;
cardiovascular disease, hypertension, cardiac arrhythmia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               68.4%; Score 93; DB 22; Length 49; 66.7%; Pred. No. 9.8e-06; ive 1; Mismatches 5; Indels
                                                                                                                                                                                                                                                 analyzing gene expression in human fetal liver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Rank DR;
                                                                                                                              Rank DR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABB22584 standard; Protein; 49 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2 НИНИНИНИНИНИН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (MOLE-) MOLECULAR DYNAMICS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     04-FEB-2000; 2000US-0180312.
26-MAY-2000; 2000US-0207456.
30-UTN-2000; 2000US-0608408.
03-AUG-2000; 2000US-023366.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0234687.
04-OCT-2000; 2000GB-0024263.
                                                                                                                              Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          30-JAN-2001; 2001WO-US00666
27-SEP-2000; 2000US-0236359.
04-OCT-2000; 2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           congenital heart disease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2001-488899/53.
                                                                                                                                 Hanzel DK,
                                                                                                                                                                                 WPI; 2001-483447/52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            49 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO200157274-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             23-JAN-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Penn SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ABB22584;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
```

ö

```
The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart (see ABA21353-ABA41305). The present sequence is a protein encoded by one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the hypertension, cardiac arrhythmias and congenital heart disease. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/published_pct_sequences.
                        Claim 15; SEQ ID No 24354; 530pp; English.
```

49 AA; Seguence

Gaps ö Query Match 68.4%; Score 93; DB 22; Length 49; Best Local Similarity 66.7%; Pred. No. 9.8e-06; Matches 12; Conservative 1; Mismatches 5; Indels

ö

ઠે 셤

Search completed: January 20, 2004, 18:23:18 Job time : 31.0145 secs

	- Alexandra Contraction and the second of th	The state of the s	, 177 7			4
3. 4						
		\$	-1-			
		, (O)				
ji.		· C.			•	* .
**	with the state of	*	-0.			
					•	
			•	•		
b .				•		•
			. *			
			.;			
	-8-1111					
3				•		
i.						
6	. ·					
N.						

						•
6		•				
	the transfer of the second sec					
1						
			٠.			
					: :	
100				•		
2.5						
1						
	* * *	•				
			•			
and sink it.						•
						:
1		9.2	w.			. P.
1			: • — ₁₀₀	•		
1					y- 1	41 41
		22.	· · · · · · · · · · · · · · · · · · ·			
	· · · · · · · · · · · · · · · · · · ·		* • •	QX 2	1	
		AN A		Apple of the second		•.
			1			two
						♥
Albert .	and the second of the second o					- 44

us-10-018-103a-5.rai

```
RESULT 1
US-09-328-352-4930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-328-352-4930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-461-325-453
                                                                                                                                                                                                                                      70.5
70.5
69
69
69
68
67.5
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 4, Appli
Sequence 2, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 160, Appli
Sequence 160, Appli
Sequence 30855, A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 25291, A
Sequence 2, Appli
Sequence 2, Appli
Sequence 6, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                           (without alignments)
74.858 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 17, Requence 4, Apsequence 16, Requence 16, Requ
                                                                                                                                                                  January 20, 2004, 18:22:04 ; Search time 11.3043 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 4
Sequence 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                      GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PCT US95-05772-1
US-09-461-325-160
US-09-252-991A-30855
US-09-996-243-95
US-08-557-006C-43
US-09-252-991A-25291
US-09-126-980-2
US-09-476-482-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-208-742-2
US-09-312-295-4
US-09-670-979-4
US-09-670-979-4
US-09-252-571-2
US-09-434-065-2
US-08-789-275-4
US-08-789-275-4
US-08-789-275-4
US-08-789-275-4
US-08-789-275-4
US-08-203-532F-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PCT-US95-01882A-2
US-08-255-457-1
US-09-115-032-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-211-942-17
US-08-203-532F-4
US-08-950-860-16
                                                                                                                                                                                                                                                                                                                                                                                                                                     328717 segs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                    OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                US-10-018-103A-5
136
1 KHKHHKHHKHKHKHK 20
                                                                                                                                                                                                                                                                                                                                                              BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Issued Patents AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82.5
82.5
79.5
79.5
79.5
79.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      77.5
77.5
77.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      73.5
73.5
72.5
72.5
                                                                                                                                                                                                                                                                                                                                                                Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   74.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             73.5
                                                                                                                                                                                                                                                                                         Perfect score:
                                                                                                                                                                                                                                                                                                                     Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                       Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Database
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Result
No.
                                                                                                                                                                    Run on
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               \frac{1}{2}
```

```
Sequence 4930, Application US/09328352

Batent No. 6562958

GENERAL INFORMATION:

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

TITLE OF INVENTION NUMBER: US/09/328,352

CURRENT FILING DATE: 1999-06-04

SEQ ID NOS: 8252

LENGTH: 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
     Appli
                                                                                                                                                                                                                                                           Sequence 2962, Ap
                                                                                                                                                                                                                                                                           Sequence 49, Appl
Sequence 6, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                          Sequence 36
                                                                                                                        Sequence 21
                                                                                      Sequence 2,
                                                                                                                                                            Sequence 3
                       Sequence
Sequence
Sequence
                                                                                                                                                                                          Sequence
Sequence
Sequence
                                                                        Sequence
                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 92; DB 4; Length 363;
Pred. No. 7.2e-05;
1; Mismatches 5; Indels
US-09-078-465-4
PCT-US95-01882A-4
US-09-686-528A-3
US-09-456-287-2
US-09-456-287-2
US-09-456-287-2
US-09-46-71-287-2
US-09-149-097D-36
US-08-149-097D-35
US-08-149-097D-35
US-08-292-2
US-08-311-644-2
PCT-US93-04102-2
US-08-314-644-2
US-08-314-644-2
US-08-34-4288-49
US-08-31-644-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAPPLICATION THE ABOUT TITLE OF INVENTION: 94 Human Secreted Proteins TITLE OF INVENTION: 94 Human Secreted Proteins FILE REFERENCE: PSO209P1
CURRENT PILING DATE: 1999-12-14
EARLIER APPLICATION NUMBER: PCT/US99/13418
EARLIER FILING DATE: 1999-06-15
EARLIER FILING DATE: 1999-06-15
EARLIER FILING DATE: 1999-06-16
EARLIER PILING DATE: 1998-06-16
EARLIER PILING DATE: 1998-06-16
EARLIER PILING DATE: 1998-06-16
EARLIER PILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/089,509
EARLIER PILING DATE: 1998-06-16
EARLIER PILING DATE: 1998-06-12
EARLIER PILING DATE: 1998-06-12
EARLIER PILING DATE: 1998-06-22
EARLIER PILING DATE: 1998-06-22
EARLIER PILING DATE: 1998-06-22
                                                                                                                                                                                                                                                                                                                                               ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Sequence 453, Application US/09461325A; Patent No. 6475753; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: PRT
ORGANISM: Acinetobacter baumannii
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     188 НИНИНИНИНИНИНИ 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2 НХНИКНИКНИКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            67.6%;
66.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 66.7x
Best Local 12; Conservative
     302
302
313
313
313
313
225
2509
2509
2509
480
707
707
125
125
189
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3;
                                                                                                                                                                                                                                                                                                                                                                                                   7;
                                                                                                                                                                                                                                                                                                                                                              60.7%; Score 82.5; DB 4; Length 1199; 55.6%; Pred. No. 0.0033; ive 1; Mismatches 4; Indels 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                58.5%; Score 79.5; DB 2; Length 763; 57.1%; Pred. No. 0.0051; cive 0; Mismatches 6; Indels
                                                             GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Kaufmann, Joerg
TITLE OF INVENTION: CIF130 INHIBITS CELL CYCLE PROGRESSION
FILE REPERENCE: 200130.456 / 1513.003
FILE REPERENCE: 200130.456 / 1513.003
CURRENT PELLING NUMBER: US/09/709,979
PRIOR APPLICATION NUMBER: US 09/332,295
PRIOR FILING DATE: 1999-06-11
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER: FIGORY:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: DEM PC-DOS/MS-DOS
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/677,862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PILING DATE:
CLASSIFICATION: 435
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Breaner, David J
REGISTRATION NUMBER: 24,774
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415,781,1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1142 KHSDHHHHHHHHHEKKKKKKKKKHKHKHKHKHK 1168
                                                                                                                                                                                                                                                                                                                                                                                                                                           1 КНКНИКНИКН-----нКН-нКНК 20
                       Sequence 4, Application US/09709979 Patent No. 6423822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRANDEDNESS: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEFAX: 415 398-3249
INFORMATION FOR SEQ ID NO: 3
SEQUENCE CHARACTERISTICS:
LENGTH: 763 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 57.1
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                           15; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            , MOLECULE TYPE: peptide US-08-677-862-2
                                                                                                                                                                                                                                                                                              TYPE: PRT
ORGANISM: Homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 15; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CA
USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 94111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 6
US-08-677-862-2
                                                                                                                                                                                                                                                                                                                                    US-09-709-979-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE:
                                                                                                                                                                                                                                                           SEQ ID NO 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7;
                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 60.7%; Score 82.5; DB 4; Length 1199; 55.6%; Pred. No. 0.0033; ive 1; Mismatches 4; Indels 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 82.5; DB 3; Length 1199;
Pred. No. 0.0033;
1; Mismatches 4; Indels 7
                                                                                                                                                                  Query Match
61.8%; Score 84; DB 4; Length 148;
Best Local Similarity 76.5%; Pred. No. 0.0003;
Matches 13; Conservative 1; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 4, Application US/09332295
Patent No. 630372
GENERAL INFORMATION:
APPLICANT: Kaufmann, Joerg
TITLE OF INVENTION: CIF130 INHIBITS CELL CYCLE PROGRESSION
FILE REFERENCE: 200130.456 / 1513.003
CURRENT APPLICATION NUMBER: US/09/332,295
CURRENT FILING DATE: 1999-06-11
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Kaufmann, Joerg
TITLE OF INVENTION: CIR150/hTAFII150 is Necessary for Cell
TITLE OF INVENTION: Cycle Progression
FILE REPRENCE: 1453.002
CURRENT APPLICATION NUMBER: US/09/208,742
CURRENT FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 КНКИНКИНКИНКИ-----НКИ-НКИК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 КНКИНКИНКИНКИ-----НКИ-НКИК 20
                                                                                                                                                                                                                                                                                                                                                                            Sequence 2, Application US/09208742
Patent No. 6174679
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                         119 KMHKHQKHHKYHKHGKH 135
EARLIER FILING DATE: 1998-06-22
NUMBER OF SEQ ID NOS: 532
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 453
LENGTH: 148
                                                                                                                                                                                                                                                 з кинкинкинки 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        55.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              , TYPE: PRT
; ORGANISM: Homo sapien
US-09-332-295-4
                                                                                                TYPE: PRT
CORGANISM: Homo sapiens
US-09-461-325-453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 15; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
Matches 15; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; TYPE: PRT
; ORGANISM: human
US-09-208-742-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEO ID NO 4
LENGTH: 1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 4
US-09-332-295-4
                                                                                                                                                                                                                                                                                                                                                                    US-09-208-742-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                        g
```

;

Gaps

```
APPLICANT: Paritoly.

APPLICANT: Prichard, Melanie
APPLICANT: Prichard, Melanie
APPLICANT: Vilano, Jozdi Guimera
TITLE OF INVENTION: Human Gene Sequence of the Down Syndrome Critical
Patent No. 6251664
TITLE OF INVENTION: Region of Human Chromosome 21, Coding for A Serine-
TITLE OF INVENTION: Threonine Protein Kinase (MMB), Expressed in the
TITLE OF INVENTION: Neuronal Regions Affected in Down Syndrome
TITLE OF INVENTION: UNMBRR: US/08/789,275A
CURRENT APPLICATION NUMBER: 1997-01-28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ب
.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ë
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 763;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ch 58.5%; Score 79.5; DB 3; Length 763; 1 Similarity 57.1%; Pred. No. 0.0051; 12; Conservative 0; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 58.5%; Score 79.5; DB 3; Length 7 Best Local Similarity 57.1%; Pred. No. 0.0051; Matches 12; Conservative 0; Mismatches 6; Indels
                                                                                                                  CURRENT APPLICATION DATA:

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/434,065

FILING DATE:

CLASSIFICATION NUMBER: US/09/434,065

FILING DATE:

ATONIN'S APPLICATION DATA:

APPLICATION NUMBER: 08/677,862

FILING DATE:

ATONIN'S APPLICATION NUMBER: 24,774

REGISTRATION NUMBER: 24,774

REGISTRATION NUMBER: 24,774

REGISTRATION NUMBER: 24,774

RELEPROCHOWINICATION INFORMATION:

TELEPROCH STATE OF 191-1989

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHRARACTERISTICS:

LEMETH: 763 amina acids
                                                       MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     599 HHHHGNSSHHHHHHHHHHH 619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         599 HHHHGNSSHHHHHHHHHHHHH 619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2 нкин---кинкинки 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 НКН---КНИКНИКНИМ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 4, Application US/08789275A; Patent No. 6251664; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 763 amino acide
TYPE: amino acid
STRANDEDNESS: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-434-065-2
                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NUMBER OF SEQ ID NOS: 6
SEQ ID NO 4
LENGTH: 763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: PRT
ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
Matches 12; Conserv
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-789-275-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-789-275-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ë,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      58.5%; Score 79.5; DB 2; Length 763; 57.1%; Pred. No. 0.0051; Live 0; Mismatches 6; Indels 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 2, Application US/09434065
; Patent No. 6107074
; GENERAL INFORMATION:
    APPLICANT: SONG, Ho Yeong
; TITLE OF INVENTION: TRAP2-A880ciated Kinase
; TUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS: 3
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
; STATE: CA
                                                                                                                                                                                                                             APPLICANT: SONG, Ho Yeong
APPLICANT: ROTHE, Mike
TITLE OF INVENTION: TRAF2-Associated Kinase
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSE: Flahr, Hobbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, 34th floor
CITY: San Francisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SURTHWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/252,571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       T96-005/A63613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CLASSIFICATION:
PRIOR APPLICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/677,862
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BESTREY. 24 774
REFERENCE/DOCKET NUMBER: 24,774
REFERENCE/DOCKET NUMBER: T96-005/
TELEPHONE: 415 781-1989
TELEPHONE: 415 781-1989
INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS:
LENGTH: 763 maino acids
TYPE: maino acids
TYPE: maino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                              ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Ploppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               599 HHHHHHHHHHHHHHHHHH 619
                                                   599 HHHHGNSSHHHHHHHHHHHHH 619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2 НКН---КНИКНИКНИКН 19
                    2 НКНИ---КНИКНИКНИКИ 19
                                                                                                                                                                   Sequence 2, Application US/09252571
Patent No. 5981250
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                     cA
USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: amino a STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                          STATE: C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 8
US-09-434-065-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-252-571-2
                                                                                                                                                 US-09-252-571-2
                                                                                                                        RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                       셤
                    8
```

```
Score 78; DB 2; Length 303;
Pred, No. 0.0032;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: PatentIn Release #1.0, Version #1.25 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/078,465 FILING DATE: 14-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 2, Application PC/TUS9501882A
GENERAL INFORMATION:
APPLICANT: Gorski, David H.
APPLICANT: Walsh, Kenneth
TITLE OF INVENTION: Growth Arrest Homeobox Gene
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calfee, Halter, and Griswold
STREET: 800 Superior Avenue
                                                                                                                                                                                                                 RESULT 12
US-078-465-2
US-078-465-2
Sequence 2, Application US/09078465
Patent No. 6280969
GENERAL INFORMATION:
APPLICANT: Walsh, Kenneth
TITLE OF INVENTION: Growth Arrest Homeobox Gene
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                     Calfee, Halter, and Griswold
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION DATA:
PRIOR APPLICATION NUMBER: US/08/203,532
APPLICATION NUMBER: US/08/203,532
ATTORNEY AGENT INFORMATION:
NAWE: GOLFICK, MARY E.
REGISTRATION NUMBER: 34829
REFERENCE/DOCKET NUMBER: 22311/00114
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
                                                      3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              з кникничикничик 20
                                                                                                з кинкинкинкникик 20
                                                                                                                                 :||: | || || || || 63 ОННВСНИННИННИНО 80
                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: Calfee, Halter, STREET: 800 Superior Avenue CITY: Cleveland
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 57.4%;
Best Local Similarity 55.6%;
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (216) 622-8200
(216) 241-0816
       Query Match 57.4%;
Best Local Similarity 55.6%;
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : 303 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ZIP: 44114-2688
COMPUTER READABLE FORM:
MEDIUM TYPE: Florm:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEX: 980499
INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS: LENGTH: 303 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-078-465-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PCT-US95-01882A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                            ઠ
                                                                                                                                                  셤
                                                                           GENERAL INCORPATION:
GENERAL INCORMATION:
GENERAL INCORMATION:
APPLICANT: Palleja, Estivill Xavier
APPLICANT: Pritchard, Melanie
APPLICANT: Pritchard, Melanie
APPLICANT: Vilaro, Jordi Guimera
TITLE OF INVENTION: Human Gene Sequence of the Down Syndrome Critical
TITLE OF INVENTION: Region of Human Chromosome 21, Coding for A Serine-
TITLE OF INVENTION: Threonine Protein Kinase (MNB), Expressed in the
TITLE OP INVENTION: Neuronal Regions Affected in Down Syndrome
TITLE OP INVENTION NUMBER: US/08/789,275A
CURRENT APPLICATION NUMBER: US/08/789,275A
CURRENT PILING DATE: 1997-01-28
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 5

LENGTH: 763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 79.5; DB 3; Length 763;
Pred. No. 0.0051;
0; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 2, Application US/08203532F
; Sequence 2, Application US/08203532F
; Patent No. 5856121
; GENERAL INFORMATION:
    APPLICANT: Gorski, David H.
; APPLICANT: Walsh, Kenneth
    TITLE OF INVENTION: Growth Arrest Homeobox Gene
    NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
    ADDRESSE: Calfee, Halter, and Griswold
    STREET: 800 Superior Avenue
    CITTY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: U.S.A.
ZIP: 44114-2688
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/203,532F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ниннсиззнинининин 619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2 НКНН----КННКННКННКН 19
                                               Sequence 5, Application US/08789275A Patent No. 6251664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: GOLTICK, MART 34829
REGISTRATION NUMBER: 34829
REFERENCE/DOCKET NUMBER: 2231
TELECHONE: (216) 622-8200
TELEPHONE: (216) 241-0816
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEX: 980499
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 57.1%;
Matches 12; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         : 303 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT
ORGANISM: Rat norvegicus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STREET: BUL
STREET: BUL
TTY: Cleveland
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-203-532F-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-203-532F-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-789-275-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 78; DB 3; Length 303;
Pred. No. 0.0032;
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                    ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2
                                                                                                                                              57.0%; Score 77.5; DB 1; Length 60; 52.2%; Pred. No. 0.0008; tive 0; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 57.0%; Score 77.5; DB 2; Length 60; Best Local Similarity 52.2%; Pred. No. 0.0008; Matches 12; Conservative 0; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Plaut, Andrew G.
APPLICANT: Gilbert-Rothstein, Joanne V.
APPLICANT: Wight, Andrew
TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL BINDING
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 3
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CITY: Boston
STATE: Massachusette
COMPUTE: Massachusette
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/09/115,032
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/255,457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATTORNEI/AGGE,

NAME: Clark, Paul C.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Search completed: January 20, 2004, 18:27:12
                                                                                                                                                                                                                       2 НИНИИНИНИН-----НИНКН 19
                                                                                                                                                                                                                                                           11 іннінжінтінжінжі 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11 нинининтининунссвинини 33
                                                                                                                                                                                                                                                                                                                                    US-09-115-032-1
; Sequence 1, Application US/09115032
Patent No. 5972348
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATTORNEY/AGENT INFORMATION:
                                    rrPE: amino acid;
; TOPOLGGY: linear;
; MOLECULE TYPE: protein
                                                                                                                                  Query Match
Best Local Similarity 52.2'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: (617) 542-890
TELEX: 200154
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             60 amino acids
                                    60 amino acids
INFORMATION FOR SEQ ID NO:
                   SEQUENCE CHARACTERISTICS LENGTH: 60 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               , MOLECULE TYPE: protein US-09-115-032-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                         ઠે
                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            57.4%; Score 78; DB 5; Length 303; 55.6%; Pred. No. 0.0032;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..0032;
... 5; Indels
                             COUNTRY: U.S.A.
ZIP: 44114-2688
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMpatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
PAPLICATION NUMBER: PCT/US95/01882A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/255,457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul C.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEPAX: (617) 542-8906
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3; Mismatches
                                                                                                                                                                                                FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: GOLIACK, MAY E.
REGIGTRATION NUMBER: 34829
REFERENCE/DOCKET NUMBER: 2231
TELEFORMUNICATION INFORMATION:
TELEFORMUNICATION 1NFORMATION:
TELEFACMUNICATION 1NFORMATION:
TELEFACM (216) 622-8200
TELEFACM (216) 241-0816
TELEFAC (216) 241-0816
TELEFAC (309) 300 100: 2: SEQUENCE CHARACTERISTICS:
LENGTH: 303 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            E: Fish & Richardson
225 Franklin Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 КНИКНИКНИКНИКИК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63 OHHRGHEITHEHO 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 55.6
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-01882A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CITY: Boston
STATE: Massachusetts.
COUNTRY: U.S.A.
Cleveland
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEFAX: (b1.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-255-457-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
```

ઠે a

Job time : 12.3043 secs

us-10-018-103a-5.rai

```
January 20, 2004, 18:23:25; Search time 22.6087 Seconds (without alignments) 180.887 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1: (cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
| cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
| cgn2_6/ptodata/1/pubpaa/NEW_PUB.pep:*
| cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
| cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
| cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
| cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
| cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
| cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
| cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
| cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
| cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
| cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                762491 segs, 204481190 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Maximum Match 100%
Listing first 45 summaries
                                                                                                                                             OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                    136
1 КНИНИНИНИНИКНИКН 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Post-processing: Minimum Match 08
                                                                                                                                                                                                                                                                                                                                                                         US-10-018-103A-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Database
                                                                                                                                                                                                                        Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Description	Sequence 5, Appli Sequence 5, Appli	Sequence 8, Appli	Sequence 15, Appl	15,	Sequence 33892, A	Sequence 37882, A	Sequence 33313, A	Sequence 31185, A	Sequence 37352, A	Seguence 39, Appl	Sequence 44, Appl	Sequence 46, Appl	Sequence 35, Appl
SUMMARIES	US-10-131-909A-5 US-10-018-103A-5	US-10-131-909A-8	US-10-131-909A-15	US-10-018-103A-15	US-10-029-386-33892	US-09-864-761-37882	US-09-864-761-33313	US-10-029-386-31185	US-09-864-761-37352	US-10-315-515-39	US-10-315-515-44	US-10-315-515-46	US-10-315-515-35
	12 51	12	12	15	12	6	σ	12	6	12	12	12	12
* Query Match Length DB	20	26	15	15	378	49	82	108	78	90	90	93	95
% Query Match	100.0	80.1	77.2	77.2	69.1	68.4	67.6	6.99	66.2	66.2	66.2	66.2	66.2
Score	136	109	105	105	94	93	92	91	90	90	90	90	90
Result No.	7	m <	'n	9	7	60	60	10	11	12	13	14	15

232222222222222222222222222222222222222	96 96 96 123 96 123 96 120 100 100 120 120 120 120 120 120 120	22 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	96 96 96 123 96 123 96 120 100 100 120 120 120 120 120 120 120	34,	Sequence 36, Appl	Sequence 37, Appl	Sequence 40, Appl	Sequence 41, Appl	Sequence 42, Appl	Seguence 45, Appl	Sequence 43, Appl	Sequence 38, Appl	7				Seguence		Sequence	Sequence	Sequence	Sequence					229 Sequence 33229, A	6		4	2	Sequence 258	10000
	96 96 96 96 100 100 128 128 128 128 138 1199 1199 1199 1199 1199 1199 1199	96 96 96 96 100 100 128 128 128 128 138 1199 1199 1199 1199 1199 1199 1199	66.2 66.2	US-10-315-515-34	US-10-315-515-36	0-315-5	0-315-515	US-10-315-515-41	US-10-315-515-42	US-10-315-515-45	US-10-315-515-43	US-10-315-515-38	US-09-864-761-3372	US-09-864-761-34744	US-09-764-868-76	US-09-864-761-379	US-10-104-307-17	US-10-029-386-30	US-10-296-144-5	US-10-012-542-45	US-10-106-698-59					US-10-338-279-4	US-10-029-386-33	US-09-864-761-416	US-10-131-909A-4	US-10-018-103A-4	US-10-211-133-7	US-10-097-340-25	70 00 LAC AND OC OT
			0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	96 12	96 12	96 12	36 12	36 12	96 12	36 12)5 12)6 12	37 9	37 9	51 10	92 9	16 12	29 12	34 15	18 15	76 15	•	•	•	•	99 15	57 12	99	29 12	29 15	15	13 15	•

ALIGNMENTS

```
US-10-131-999A-5;
Sequence 5, Application US/10131909A;
Sequence 5, Application Wo. US20030165567A1;
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL INFORMATION:
Histidian Copolymer and Methods For Using Same;
TITLE OF INVENTION: Histidian Copolymer and Methods For Using Same;
TITLE OF INVENTION: Histidian Copolymer and Methods For Using Same;
FILE REFERENCE: 5627*6
CURRENT FILING DATE: 2001-11-05;
PRIOR APPLICATION NUMBER: US 10/018103
PRIOR FILING DATE: 2000-12-20
PRIOR PRIOR APPLICATION NUMBER: US 60/173576
PRIOR FILING DATE: 1999-12-29;
NUMBER OF SEQ ID NOS: 17
SEQ ID NO 5
LENGTH: 20
TYPE: RRT
ORGANISM: Artificial Sequence
FEATURE:
COTHER INFORMATION: Synthetic Peptide
COTHER INFORMATION: Synthetic Peptide
COURTY MATCH
BEST LOCAL Similarity 100.0%; SCORE 136; DB 12;
BEST LOCAL Similarity 100.0%; Pred. No. 6.56-09;
MATCHES 20; CONSETVATINE 20

MATCHES 20; CONSETVATINE 20
```

Db 1 KHKHHKHKHKHKHKHK 20

ö

RESULT 2 US-10-018-103A-5 ; Sequence 5, Application US/10018103A

```
Best Local Similarity
Matches 15; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 6
US-10-018-103A-15
                                                                                                                                                                                                                                                                                                                    US-10-018-103A-8
                                                                                                                                                                                                       LENGTH: 26
                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 8, Application US/10131909A
; Sequence 8, Application US/10131909A
; Publication No. US20030165567A1
; GENERAL INFORMATION:
; APPLICAT: Mixson, A. James
; TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
; TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
; TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
; FILE REFERENCE: 5627*6
; CURRENT APPLICATION NUMBER: US/10/131,909A
; PRIOR FILING DATE: 2001-11-05
; PRIOR FILING DATE: 2001-11-20
; PRIOR FILING DATE: 2000-12-20
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: US 60/173576
; PRIOR FILING DATE: 1999-12-29
Publication No. US20030045465Al
GENERAL INFORMATION:
APPLICANT: Mixson, A. James
TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
FILE REFERENCE: 5627*5
CURRENT APPLICATION NUMBER: US/10/018,103A
CURRENT FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: PCT/US00/34603
PRIOR PELING DATE: 1999-12-29
PRIOR PELING DATE: 2000-12-20
PRIOR PILING DATE: 2000-12-20
NUMBER OF SEQ ID NOS: 16
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin version 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: Mixson, A. James
TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
FILE REFERENCE: 5627*5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 109; DB 12; Length 26;
Pred. No. 6.4e-06;
2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     , DB 15;
6.5e-09;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
100.0%; Score 136; Dest Local Similarity 100.0%; Pred. No. 6.5 Matches 20; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     , OTHER INFORMATION: Synthetic Peptide US-10-131-909A-8
                                                                                                                                                                                                                                                                                                                                                                                                            , OTHER INFORMATION: Synthetic Peptide US-10-018-103A-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 8, Application US/10018103A Publication No. US20030045465A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 КНКНИКНИКНИКНКИК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 КНКНИКНИКНИКНИК 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 83.3%;
Matches 15; Conservative ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin version 3.1
SEQ ID NO 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                           ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-10-018-103A-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-10-131-909A-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                               SEQ ID NO 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
```

```
US-10-131-909A-15

| Sequence 15, Application US/10131909A |
| Publication No. US2030165567A1 |
| Publication No. US2030165567A1 |
| GENERAL INFORMATION: Histidine Copolymer and Methods For Using Same |
| TITLE OF INVERTION: Histidine Copolymer and Methods For Using Same |
| TITLE OF INVERTION: Histidine Copolymer and Methods For Using Same |
| FILE REFERENCE: 5627*6 |
| CURRENT APPLICATION NUMBER: US/10/131,909A |
| FILE REPERSONG: 5627*6 |
| PRIOR APPLICATION NUMBER: PCT/US00/34603 |
| PRIOR FILING DATE: 2000-11-05 |
| PRIOR FILING DATE: 2000-12-29 |
| NUMBER OF FILING DATE: 1999-12-29 |
| NUMBER OF SEQ ID NOS: 17 |
| SOFTWARE: Patentin version 3.1 |
| TENDAL PATENTIAL PATENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Mixeon, A. James
TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
TITLE SETAMENCE: 522*5
CURRENT APPLICATION NUMBER: US/10/018,103A
CURRENT FILING DATE: 2001-11-05
PRIOR FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: PCT/US00/34603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           77.2%; Score 105; DB 12;
100.0%; Pred. No. 1.1e-05;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 80.1%; Score 109; DB 15;
illarity 83.3%; Pred. No. 6.4e-06;
Conservative 2; Mismatches 1;
CURRENT APPLICATION NUMBER: US/10/018,103A;
CURRENT FILING DATE: 2001-11-05;
PRIOR APPLICATION NUMBER: US 60/173576;
PRIOR FILING DATE: 1999-12-29;
PRIOR PLILING DATE: 1999-12-29;
PRIOR FILING DATE: 2000-12-20;
NUMBER: OF SEQ ID NOS: 16;
SEQ ID NO 8;
SEQ ID NO 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; OTHER INFORMATION: Synthetic Peptide
US-10-131-909A-15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 15, Application US/10018103A publication No. US20030045465A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OTHER INFORMATION: Synthetic Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9 RRRHHKHHKHHKHHK 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 КНКНИКНИКНИКНКНК 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 ННИНИНИНИНКНИК 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 77.2%
Best Local Similarity 100.0
Matches 15; Conservative
```

ö

ö

```
Sequence 3313, Application US/09864761
Parent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Rank, David R.
APPLICANT: Rank, David R.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Acomica-X-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HELLOO, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN FITAL LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN FORM MARROW, SIGNAL = 1.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 37882
LENGTH: 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 93; DB 9; 1 Pred. No. 0.00057;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1; Mismatches
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR PILING DATE: 2000-08-03
PRIOR PLLING DATE: 2000-10-04
PRIOR PLLING DATE: 2000-10-04
PRIOR PLLING DATE: 2000-10-04
PRIOR PLLING DATE: 2000-09-27
PRIOR PLLING DATE: 2000-09-27
PRIOR PLLING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/USO1/00669
PRIOR PLLING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/USO1/00669
PRIOR PLLING DATE: 2001-01-30
PRIOR PLLING DATE: 2000-09-21
PRIOR PLLING DATE: 2000-09-31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OTHER INFORMATION: MAP TO AC011186.2
OTHER INFORMATION: EXPRESSED IN PLACE
OTHER INFORMATION: EXPRESSED IN HELA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 нинининининини 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 66.7%;
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: PRT
ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-864-761-37882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-864-761-33313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 33892, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David R.
APPLICANT: Hanzel, David R.
APPLICANT: HANZEL, DAVID RENOWED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR (TITLE OF INVENTION: EXPRESSION ANALYSIS TWO FILE REFERENCE: AEOMICA-X-2 CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                69.1%; Score 94; DB 12; Length 378; 63.2%; Pred. No. 0.0026; vative 2; Mismatches 5; Indels .
                                                                                                                                                                                                                                                                                                     Length 15;
                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OTHER INFORMATION: MAP TO AD001751.1

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.6

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
US-10-029-386-33892
                                                                                                                                                                                                                                                                                             Query Match
77.2%; Score 105; DB 15;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 15; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                ; OTHER INFORMATION: Synthetic Peptide US-10-018-103A-15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         140 HEMERHEMERHEMEN 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 НКНИКНИКНИКНКИК 20
     PRIOR FILING DATE: 2000-12-20
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
                                                                                                                                      TYPE: PRT ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                        4 НИКНИКНИКНИКИ 18
                                                                                                                                                                                                                                                                                                                                                                                                                                      1 HHGHIGHIGHKHK 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT
ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-10-029-386-33892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -09-864-761-37882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQ ID NO 33892
LENGTH: 378
                                                                                  SEQ ID NO 15
                                                                                                                                                                                          FEATURE:
                                                                                                                LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
```

```
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR (TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR INVENTION: EXPRESSION ANALYSIS TWO FILE REFERENCE: AEOMICAX: US/10/029,386 CURRENT APPLICATION NUMBER: US/10/029,386 CURRENT FILING DATE: 2001-12-20 NUMBER OF SEQ ID NOS: 34288 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1 SEQ ID NO 31185 LENGTH: 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 37352, Application US/09864761.

Sequence 37352, Application US/09864761.

Baten No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Rank, David R.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 91; DB 12; Length 108;
Pred. No. 0.0019;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CTHER INFORMATION: MAP TO ACO10219.4

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3

OTHER INFORMATION: EXPRESSED IN LIVE, SIGNAL = 3.9

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.0

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6

US-10-029-386-31185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TILLE OF LINGBAILON: GENDE GATTEGSSLOW ANTALIST
CURRENT FILING DATE: 2001-06-23
PRIOR APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR PILING DATE: 2000-05-26
PRIOR PILING DATE: 2000-06-26
PRIOR APPLICATION NUMBER: US 60/23,366
PRIOR PILING DATE: 2000-09-03
PRIOR PILING DATE: 2000-09-09
PRIOR PILING DATE: 2000-09-09
PRIOR PILING DATE: 2000-09-27
PRIOR PILING DATE: 2001-01-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 НКНИКНИКНИКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      73 ненинининин 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 66.9%;
Best Local Similarity 66.7%;
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CTHER INFORMATION: EXPRESSED IN HELLIOO, SIGNAL = 8.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 9.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.4
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.0
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     67.6%; Score 92; DB 9; Length 82; 66.7%; Pred. No. 0.0011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 33313
                               CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR PILING DATE: 2000-02-04
PRIOR PELING DATE: 2000-02-06
PRIOR PELING DATE: 2000-05-26
PRIOR PLING DATE: 2000-08-03
PRIOR PLING DATE: 2000-08-03
PRIOR PLING DATE: 2000-09-03
PRIOR PLING DATE: 2000-09-03
PRIOR PLING DATE: 2000-09-27
PRIOR PLING DATE: 2000-09-27
PRIOR PLING DATE: 2000-09-27
PRIOR PLING DATE: 2001-01-30
PRIOR PRIOR DATE: 2001-01-30
PRIOR PRIOR DATE: 2001-01-30
PRIOR PLING DATE: 2001-01-30
PRIOR PELING DATE: 2001-01-30
PRIOR PELING DATE: 2001-01-30
PRIOR PELING DATE: 2001-01-30
PRIOR PELING DATE: 2001-01-30
PRIOR PLING DATE: 2000-09-21
PRIOR PLING DATE: 2000-09-21
PRIOR PLING DATE: 2000-09-30
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2 НИНИНИНИНИНИНИ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 66.7
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-864-761-33313
```

Sequence 31185, Application US/10029386 Publication No. US20030194704A1 GENERAL INFORMATION:

RESULT 10 US-10-029-386-31185

65 НИННИННИННИКИ 82

g ठ

```
Query Match 66.2%; Score 90; DB 12; Length 93; Best Local Similarity 66.7%; Pred. No. 0.0021; Matches 12; Conservative 0; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          66.2%; Score 90; DB 12; Length 90; 66.7%; Pred. No. 0.002;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6; Indels
                                                                          GENERAL INFORMATION:
APPLICANT: Wright, David A.
APPLICANT: Woytas, Daniel F.
TITLE OF INVENTION: NUCLEIC ACIDS RELATED TO PLANT
TITLE OF INVENTION: NUCLEIC MANA
TITLE OF INVENTION: RETROCLEMENTS
FILE REPERENCE: 08411-031001
CURRENT APPLICATION NUMBER: US/10/315,515
CURRENT FILING DATE: 2002-12-10
PRIOR FILING DATE: 2001-12-10
PRIOR FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 166
SOFTWARE: FREESEQ for Windows Version 4.0
SEQ ID NO 44
LENGTH: 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Wright, David A.
APPLICANT: Wright, David A.
APPLICANT: Voytas, Daniel F.
TITLE OF INVENTION: NUCLEIC ACIDS RELATED TO PLANT
TITLE OF INVENTION: RTEROCLEMENTS
FILE REFERENCE: 08411-031001
CURRENT APPLICATION NUMBER: US/10/315,515
CURRENT FILING DATE: 2002-12-10
PRIOR FILING DATE: 2001-12-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: Wright, David A.
APPLICANT: Wright, Daniel F.
TITLE OF INVENTION: NUCLBIC ACIDS RELATED TO PLANT
TITLE OF INVENTION: RETROGLEMENTS
FILE REFERENCE: 08411-031001
CURRENT APPLICATION NUMBER: US/10/315,515
CURRENT FILING DATE: 2002-12-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQ ID NOS: 168
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 46
LENGTH: 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-10-315-515-35
Sequence 35, Application US/10315515
Publication No. US/20030166190A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; Sequence 46, Application US/10315515; Publication No. US20030166190A1; GENERAL INFORMATION:
             Sequence 44, Application US/10315515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      43 нининининининин 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 ИКНИКНИНИНИНИН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 37 НЕНЕНЕНЕНЕНЕНЕН 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2 НХНИКНИНКНИКНИИ 19
                                     Publication No. US20030166190A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 66.79
Watches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: PRT
ORGANISM: Oryza sativa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM: Glycine max
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-10-315-515-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-10-315-515-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-10-315-515-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OTHER INFORMATION: MAP TO AF045448.1

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.5

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.5

OTHER INFORMATION: EXPRESSED IN HELLOO, SIGNAL = 4.5

OTHER INFORMATION: EXPRESSED IN HELLOO, SIGNAL = 3.3

OTHER INFORMATION: EXPRESSED IN HELLO, SIGNAL = 5.3

OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 5.3

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.7

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.7

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.7

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 90; DB 12; Length 90; Pred. No. 0.002; 0; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 90; DB 9; Lengtn /o;
Pred. No. 0.0018;
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/USO1/00661
PRIOR APPLICATION NUMBER: PCT/USO1/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/USO1/00670
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR FILING DATE: 2000-06-30
PRIOR FILING DATE: 2000-06-30
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 37352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 12
US-10-315-515-39
Sequence 39, Application US/10315515
Publication No. US20030166190A1
GENERAL INFORMATION:
APPLICANT: Wight, David A.
TITLE OF INVENTION: REFROELEMENTS
TITLE OF INVENTION: NUCLEIC ACIDS RELATED TO PLANT
TITLE OF INVENTION: REFROELEMENTS
FILE REFERENCE: 08411-031001
CURRENT FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: US 60/339,060
PRIOR PLILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 168
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 39
LENGTH: 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 НИНИНИНИНИНИНИН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              37 ИНИКИНИНИНИНИНИНИ 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2 НОНИМИНИНИНИНИ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    44 НИНИННИНИНИН 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 66.7%;
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 66.7%;
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; TYPE: PRT
; ORGANISM: Glycine max
US-10-315-515-39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-864-761-37352
```

ò

Gapa

ö

PRIOR APPLICATION NUMBER: US 60/339,060

US-10-315-515-44

요

ઠે

ö

Gapa

ö

us-10-018-103a-5.rapb

```
Wed Jan 21 11:27:51 2004
```

```
ö
                                                                                                                                                                                                             0; Gaps
                                                                                                                                                                 Query Match 66.2%; Score 90; DB 12; Length 95; Best Local Similarity 66.7%; Pred. No. 0.0021; Matches 12; Conservative 0; Mismatches 6; Indels
PRIOR FILING DATE: 2001-12-10

NUMBER OF SEQ ID NOS: 168

SOFTWARE: PastSEQ for Windows Version 4.0

SEQ ID NO 35

LENGTH: 95

TYPE: PR: 7

ORGANISM: Arabidopsis thaliana
US-10-315-515-35
                                                                                                                                                                                                                                                                                                                                                Search completed: January 20, 2004, 18:28:43 Job time : 23.6087 secs
                                                                                                                                                                                                                                                     2 икинкикикинки 19
| | | | | | | | | | 6
43 ининининининин 60
                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                           g
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

model	
3	
using	
search,	
protein	
•	
protein	
õ	

January 20, 2004, 18:18:04 ; Search time 10.1449 Seconds (without alignments) 189.590 Million cell updates/sec Run on:

Title: US-10-018-103A-5
Perfect score: 136
Sequence: 1 KHKHHKHKHKHKHK 20

Scoring table:

BLOSUM62 Gapop 10.0 , Gapext 0.5

283308 seqs, 96168682 residues

Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR_76:* Database :

1: pirl:* 2: pir2:* 3: pir3:* 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description	probable metal tra	hypothetical prote	cal	bable pect	sal homeotic prote	homeotic	sal homeotic prote	Ü	shs protein - frui	knob-associated hi	hypothetical prote	hypothetical prote	histidine-rich pro	probable zinc tran	histidine-rich gly	HNF-3/fork head fa	transforming prote	hypothetical prote	selenoprotein P pr	Down-syndrome-crit	conserved hypothet	hypothetical prote	homeotic protein M	homeobox protein M	growth arrest-spec	polycomb (Pc) prot	neuron-derived rec		neuron derived orp
	ΩI	4	T13389	A44971	G84783	B33910	800262	C33910	T29092	T13804	A29454	T34520	H86187	A29653	A84696	KGZOHL	A47446	A55909	T04219	OMRTSP	JC4898	B64421	AB2396	A56837	B49122	A48130	A38565	866671	O)	JC2493
	8	7	N	7	~	~	~	~	~	~	~	~	N	~	~	Н	~	~	~	Н	~	-4	7	~	~	~	н	~	~	7
	Match Length	326	1291	96	407	139	142	142	1040	1212	657	169	391	82	385	351	387	451	658	385	754	143	353	303	303	303	390	430	625	628
overy	Match	67.6	6.99	66.2		65.4	65.4	65.4	64.0	64.0	62.5	61.8	61.0	60.7	60.7	60.3	59.9	59.9	59.6	58.5	•	58.1	58.1	57.4	57.4	57.4	57.4	57.4	57.4	57.4
	Score	92	91	90	90	89	88	89	87	87	82	84	83	82.5	82.5	82	81.5	81.5	81	79.5	79.5	79	79	78	78	78	78	78	78	78
Result	No.	7	7	e	4	S	9	7	80	o,	10			13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29

probable histidine hypothetical prote	hypothetical prote hypothetical prote hypothetical prote	eggshell protein – homeotic protein H hypothetical prote	hypothetical prote hypothetical prote neuroblast prolife	URBS1 protein – sm hypothetical prote hypothetical prote	HNF-3beta - Africa hypothetical prote
C64698 H90992	T39712 T42516 T16435	A54530 WJFFH2 T00677	T46024 C85838 A40721	S27473 A64978 T20270	I51436 T50609
~~	000	2 11 22	0 0 0	~ ~ ~	77
279	420 427 83	149 419 529	580 283 474	950 274 351	434 529
57.0	56.6 55.9	55.9 55.9 9.0 9.0	55.3 55.5 5.5	55.5 55.1 55.1	55.1 55.1
77.5	77 76 76	76 76 76	76 75.5 75.5	75.5 75 75	75 75
30	w w w w c 4	35 37	8 8 8 0 0	4 4 4 1 2 6	44 45

ALIGNMENTS

RESULT 1

D83483
probable metal transporter PA1297 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 06-Jan-2003
C;Accession: D83483
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A; Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: D83483
A;Status: preliminary
A; Molecule type: DNA
A;Residues: 1-326 <sto></sto>
A;Cross-references: GB:AE004559; GB:AE004091; NID:g9947228; PIDN:AAG04686.1; GSPDB:GN00
A;Experimental source: strain PAO1

C,Genetics: A,Gene: PA1297 C,Superfamily: zinc transporter ZnT-2

Gaps ; 0 Query Match 67.6%; Score 92; DB 2; Length 326; Best Local Similarity 63.2%; Pred. No. 0.00037; Matches 12; Conservative 1; Mismatches 6; Indels

ö

| || || || || || || || 1 163 ининининини || 181 2 НКНИКНИКНИКНКИК 20 ઠે 셤

ö

hypothetical protein 115C2.10 - fruit fly (Drosophila melanogaster)
C; Species: Drosophila melanogaster
C; Species: Drosophila melanogaster
C; Date: 13.402-1999 #sequence_revision 13.40g-1999 #text_change 17-Nov-2000
C; Accession: T13389
R; Salles, C.; Valenti, P.; Darlamitsou, A.; Henderson, N.; Campbell, L.; Glover, D. R; Salles, C.; Valenti, P.; Darlamitsou, A.; Henderson, N.; Campbell, L.; Glover, D. A; Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A; Reference number: Z17665
A; Reference number: Z17665
A; Residues: T13389
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Residues: 1.1291 < CAT>
A; Residues: 1.1291 < CAT>
A; Residues: 1.1291 < CAT>
A; Cross-references: EMBL:AL031581; NID:e1320978; PID:e1426292; PIDN:CAA20894.2
C; Gentics:
A; Map position: X
A; Introns: Z38/3; 1225/1
A; Note: EG:115C2.10

ö

```
halternate names: homeotic protein = fruit fly (Drosophila melanogaster)
Nalternate names: homeotic protein spalt
C;Alternate names: homeotic protein spalt
C;Species: Drosophila melanogaster
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 20-Aug-1999
C;Accession: S00262; A38026; A33910
C;Accession: S00262; A38026; A33910
B;Prei, E: Schuh, R.; Baumgartner, S.; Burri, M.; Noll, M.; Juergens, G.; Seifert, B.;
ENBO J. 7, 197-204, 1988
A;Title: Molecular characterization of spalt, a homeotic gene required for head and tain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sal homeotic protein - fruit fly (Drosophila orena)

Gispecies: Drosophila orena
Cispecies: Drosophila orena
Cispecies: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 21-Feb-1997
Cispecession: C33910
Cispecession: C33910
Cispecession: Caluh, R.; Jaeckle, H.
Rigeuter, D.; Schuh, R.; Jaeckle, H.
Proc. Natl. Acad. Sci. U.S.A. 86, 5483-5486, 1989
A;Title: The homeotic gene spalt (8a1) evolved during Drosophila speciation.
A;Teference number: A33910; MUID:89315821; PMID:2568636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AyMolecule type: mRNA
A,Residues: 1-142 <FR2>
R,Residues: 1-142 <RR2>
R,Residues: 1-142 <RR3>
R,Residu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                           A;Residues: 1-139 <REU>
A;Cross-references: GB:M21227; NID:g158375; PIDN:AAA28877.1; PID:g158376
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Molecule type: DNA
A;Reaidudes: 1-12 <FR1>
A;Crosa-references: EMBL:X57474; NID:g8536; PIDN:CAA40712.1; PID:g8537
A;Accession: A38026...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 142;
                                                                                                                                                                                                                                                                                                                                                                                                                             Length 139;
A;Reference number: A33910; MUID:89315821; PMID:2568636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0.00038;
                                                                                                                                                                                                                                                                                                                                                                                                                 / Match 65.4%; Score 89; DB 2; I Local Similarity 64.7%; Pred. No. 0.00038; neg 11; Conservative 4; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  65.4%; Score 89; DB 64.7%; Pred. No. 0.00:ive 4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   F.19-80/Region: glycine-rich
F.92-99,103-110/Region: 8-residue repeats
F.108-142/Region: histidine-rich
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A, Cross-references: FlyBase: FBgn0003313
A, Map position: 2L 33A1,2
                                                                                                                                                                                                                                                                      A,Gene: FlyBage:Dsim/sala
A,Cross-references: FlyBase:FBgn0012892
C,Superfamily: homeotic protein sal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     :||:|| ||:|
126 EHHEHHGHHEHHGHRH 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C; Superfamily: homeotic protein sal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 КНИКНИКНИКНИКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :||:|| ||:|| ||:|
123 ЕННЕННGНHEHHGHHRH 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 КНИКНИКНИКНИМИКИ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 11; Conserv
                                               A, Accession: B33910
A, Status: preliminary
A, Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Accession: S00262
                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Rilin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.; Rilin, X.; Kaul, M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID: 20083487; PMID: 10617197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C;Accession: A44971
R;Kilejian, A.; Yang, Y.F.; Cochrane, A.H.; Rashid, M.A.
Mol. Biocham. Parasitol. 38, 291-293, 1990
A;Title: Homologous sequences in Plasmodium cynomolgi and the gene of the histidine-rich A;Reference number: A44971; MUID:90220761; PMID:2183051
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A, Cross-references: GB: AE002093; NID: 94415916; PIDN: AAD20147.1; GSPDB: GN00139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ### said to be proted in - fruit fly (Drosophila simulans)

C; Species: Drosophila simulans

C; Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 20-Aug-1999

C; Accession: B3331,

R; Reuter, D; Schuh, R.; Jaeckle, H.

P; Reuter, D; Schuh, Sci. U.S.A. 86, 5483-5486, 1989

Proc. Natl. Acad. Sci. U.S.A. 86, 5483-5486 during Drosophila speciation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          probable pectinesterase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Peb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: G84783
                                                                                                                                                                                                                                                                                                                                                                                                                                     hypothetical protein 1 - Plasmodium brasilianum
C:Species: Plasmodium brasilianum
C:Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 17-Feb-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 407;
                                   Length 1291;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 90; DB 2; Length 407
Pred. No. 0.00076;
1; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 66.2%; Score 90; DB 2; Length 96; Best Local Similarity 61.1%; Pred. No. 0.00021; Matches 11; Conservative 1; Mismatches 6; Indels
                                                                                                           5; Indels
                           Query Match
66.9%; Score 91; DB 2;
Best Local Similarity 66.7%; Pred. No. 0.0017;
Matches 12; Conservative 1; Mismatches 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 НИНКНИКНИКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        50 нкининининин 67
                                                                                                                                                                                                                                              |:|| || || || || 639 НОННЕННИННИННЕ 556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 НКНИКНИКНИКНИКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        56 HSYHSHHSHHSHHSH 73
                                                                                                                                                                                          2 НХНИКНИКНИКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  66.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A,Status: preliminary
A,Molecule type: DNA
Residues: 1-96 «KIL»
A,Cross-references: GB:M28064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity 66.7
12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-407 <STO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Accession: A44971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A, Gene: At2g36710
A, Map position: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local S:
Matches 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                      δ
```

ö

ò 셤

```
C;Species: Plasmodium falciparum
C;Species: Plasmodium falciparum
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jun-2000
C;Accession: A29454,
R;Triglia, T: Stahl, H.D.; Crewther, P.E.; Scanlon, D.; Brown, G.V.; Anders, R.P.; Kem, EMBO J. 6, 1413-1419, 1987
A;Title: The complete sequence of the gene for the knob-associated histidine-rich prote A;Reference number: A29454, MUID:87275836; PMID:3301325
A;Molecule type: DNA
A;Molecule type: DNA
A;Residues: 1-657 cTRI>
A;Cross-references: GB:Y00060; NID:g9908; PIDN:CAA68268.1; PID:g9909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ritheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K. ansen, N.F.; Hughes, B.; Huizar, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K. Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, A;Title: Sequence and analysis of Chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
                          knob-associated histidine-rich protein precursor - malaria parasite (Plasmodium falcipa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hypothetical protein DKFzp564J157.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Species: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T34520
R;Bloccker, H; Boecher, M; Brandt, P; Mewes, H.W.; Gassenhuber, J.; Wiemann, submitted to the Protein Sequence Database, October 1999
A;Reference number: Z21539
A;Accession: T34520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Cross-references: GB: AE005172; NID: 92388565; PIDN: AAB71446.1; GSPDB: GN00141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 21-Jun-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61.8%; Score 84; DB 2; Length 169; 76.5%; Pred. No. 0.0017; ive 1; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 657;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Cross-references: EMBL;AL122042
A;Experimental source: fetal brain; clone DKFZp564J157
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 85; DB 2;
Pred. No. 0.0044;
3; Mismatches 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 59 К<u>Q</u>НЕННИНИНИНИОНО 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      140 KMHKHQKHHKYHKHGKH 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 КНИКНИКНИКНИКНКИК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3 КНИКНИКНИКИ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 62.5%;
Best Local Similarity 61.1%;
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A Status: preliminary
A Molecule type: mRNA
A, Residues: 1-169 <BLO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
Matches 13; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-391 <STO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Note: DKFZp564J157.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                She protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C;Accession: T13804
R;Treisman, J.E.; Lai, Z.C.; Rubin, G.M.
Bovelopment 121, 2835-2845, 1995
A;Title: Shortsighted acts in the decapentaplegic pathway in Drosophila eye development A;Reference number: Z17767; MUID:96038094; PMID:7555710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Reaidues: 1-1040 <SAW>
A;Cross-references: EMBL:D82364; NID:d1067801; PID:d1012232; PIDN:BAA11565.1
A;Experimental source: lens fibers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TSC-22 protein homolog - chicken
C;Species: Gallus gallus (chicken)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molocule type: mRNA
A;Residues: 1-1212 <TRE>
A;Cross-references: EMBL:L42512; NID:g833709; PID:g833710; PIDN:AAC41608.1
                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R;Sawada, K.; Agata, K.; Eguchi, G. submitted to the EMBL Data Library, December 1995
A;Description: Analysis of the cDNA library of chicken lens fibers:.
A;Reference number: 220561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
64.0%; Score 87; DB 2; Length 1040;
Best Local Similarity 66.7%; Pred. No. 0.0039;
Matches 12; Conservative 0; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 1212;
                                                                                                                                                                                                                                                                                          Length 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 64.0%; Score 87; DB 2; Length 121
Best Local Similarity 64.7%; Pred. No. 0.0045;
Matches 11; Conservative 2; Mismatches 4; Indels
                                                                                                                                                                                                                                                                                   Query Match 65.4%; Score 89; DB 2; I Best Local Similarity 64.7%; Pred. No. 0.00038; Matches 11; Conservative 4; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A,Accession: T29092
A,Status: preliminary; translated from GB/EMBL/DDBJ
A,Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Cross-references: FlyBase: FBgn0010460
                                                                                                                                                   A;Gene: FlyBase:Dore/sala
A;Cross-references: FlyBase:FBgn0012624
C;Superfamily: homeotic protein sal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         :||:|| ||:|| ||:|
126 EHHEHHGHHEHHGHRH 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2 НИНИИНИНИНИН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                306 ННННИНННННЫНО 322
                                                                                                                                                                                                                                                                                                                                                                                                                      3 КНИКНИКНИКНИКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4 НИКНИКНИКНИКИК 20
                                                                                       A; Cross-references: GB: M21579
                              A;Molecule type: DNA
A;Residues: 1-142 <REU>
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Accession: T29092
                                                                                                                            C, Genetics:
```

8 g RESULT 10

a

ઠે

ŝ

ö

ö

```
R;Ravetch, J.V.; Feder, R.; Pavlovec, A.; Blobel, G.
Nature 312, 616-620, 1984.
A;Title: Primary structure and genomic organization of the histidine-rich protein of the A;Reference number: A22692; MUID:85061618; PMID:6095114
                                                                                                                                                                    A;Molecule type: DNA
A;Resdues: 1-351 - RAA>
A;Crosd-references: GB:X01469; NID:99997; PIDN:CAA25698.1; PID:99999
C;Comment: There are two copies of 16-residue repeats, two copies of 17-residue repeats
                                                                                                                                                                                                                                                                                                A)Introns: 23/3
C;Superfamily: plasmodium histidine-rich protein
C;Superfamily: plasmodium histidine-rich protein
C;Superfamily: plasmodium histidine-rich predicted (SIG)
C;Superds: glycoprotein; tandem repeat
F;1-23/Domain: propeptide #status predicted (*RO)
F;24-47/Domain: propeptide #status predicted (*RO)
F;25-74,75-90/Region: 16-residue repeats
F;124-138,139-153/Region: 17-residue repeats
F;124-138,139-153/Region: 10-residue repeats
F;123-301,312-331/Region: 10-residue repeats
F;40/Binding site: carbohydrate (Asn) (covalent) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 82; DB 1; Length 351
Pred. No. 0.0055;
0; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Search completed: January 20, 2004, 18:24:44 Job time : 12.1449 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 215 нинининин 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  60.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 НИМНИМНИМНИМ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 68.8
Matches 11; Conservative
                                                                                                                                             A;Accession: A22692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                          Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Richn, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Richn, X.; Kaul, M.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraeer, C.M.; Venter, J. Nature 402, 761-768, 1999

Nature 402, 761-768, 1999

Aritle: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A.Reference number: A84420; MuID: 20083487; PMID: 10617197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C; Species: Plasmodium falciparum
C; Species: 31-Dec.1988 #sequence_revision 31-Dec-1988 #text_change 09-Jun-2000
C; Accession. A29653
R; Lensera, R.; d'Auriol, L.; Andrieu, B.; Le Bras, J.; Galibert, F.
Biochem. Biophys. Res. Commun. 146, 368-377, 1987
A; Title: Cloning and sequencing of Plasmodium falciparum DNA fragments containing repeti
A; Reference number: A90134; MUID:87270765; PMID:3038111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ä
                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; cross references: GB: AE002093; NID: g3980394; PIDN: AAC95197.1; GSPDB: GN00139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         histidine-rich glycoprotein precursor - Plasmodium lophurae
C;Species: Plasmodium lophurae
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C;Accession: A22692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      probable zinc transporter [imported] - Arabidopsis thallana C;Species: Arabidopsis thallana (mouse-ear cress) C;Species: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001 C;Accession: A84696
                                                                                                                                                                                                                                                                                                                                                                                                                                                         letidine-rich protein - malaria parasite (Plasmodium falciparum)
Species: Plasmodium falciparum
Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jun-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1,
                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           60.7%; Score 82.5; DB 2; Length 385; 59.1%; Pred. No. 0.0052; ive 1; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 82.5; DB 2; Length 82;
Pred. No. 0.0013;
                                                                                                                                                61.0%; Score 83; DB 2; Length 391;
68.8%; Pred. No. 0.0047;
tive 1; Mismatches 4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Molecule type: DNA
A;Residues: 1-82 <LEN>
A;Cross-references: GB:M17028; NID:g160339; PID:g160341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 НКНИКНИКНИКН 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 НКНИКНИКНИКНКИК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |:|| ||:|| ||:||:
22 НКНН-ННКНИНКНИНКНО 39
                                                                                     A,Map position: 1
C,Superfamily: pectinesterase pemB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 60.7%;
Best Local Similarity 57.9%;
Matches 11; Conservative '
                                                                                                                                                                                                                                                                           5 НКНИКНИКНКИКИК 20
                                                                                                                                                                                                                                                                                                                                47 інніннінніннік 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13; Conservative
                                                                                                                                                              Query Match
Best Local Similarity 68.8<sup>†</sup>
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 13; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A Status: preliminary
A Molecule type: DNA
A Residues: 1-385 <STO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A, Gene: At2g29410
A, Map position: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C;Genetics:
                                                          C, Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                         염
```

ö

Gaps

ö

Length 351;

н

us-10-018-103a-5.rsp

```
STANDARD;
628
59
420
1158
1159
1193
1195
1195
1195
410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; B33910; B33910.
                                                                                                                                                                                                                                                                                                                                                                                                SALA DROSI
P21749;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SALA DROME
P21750;
78
77.5
77.7
76.5
76.5
76
76
76
76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SALA_DROME
ID SALA DF
DT 01-MAY.
DT 01-MAY.
DT 01-MAY.
DT 01-MAY.
DT 01-MAY.
OC SALA OI
OC BURATY
                                                                                                                                                                                                                                                                                                                                                                                SALA_DROSI
                                                                                                                                                                                                                                                                                                                                                           RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 2
  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     drosophila
plasmodium
mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             avian sarco
gallus gall
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               homo sapien
mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     rattus norv
rattus norv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              didelphis m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           brachydanio
brachydanio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          rattus norv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               methanococc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          rattus norv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P21749 drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              drosophila
drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           plasmodium
                                                                                                                                          January 20, 2004, 18:06:29; Search time 6.66667 Seconds (without alignments) 141.080 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P21748

Q2423

Q64733

Q64733

Q64733

Q61352

Q90564

Q90956

Q90956

Q90956

Q90956

Q90956

Q9095

Q6132

Q6132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P39020
P26017
Q63516
P56224
Q90436
Q92570
                                        Compugen Ltd
                     GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compuç
                                                                                                                                                                                                                                                                                                                                                                                                                   Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                            127863 segs, 47026705 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SALA_DROME
SALA_DROME
BUNZ_DROME
BUNZ_DROME
RNOB_PLAFN
FXR2_MOUSE
HRP3_PLAFS
HRP3_PLAFS
HRP3_PLAFS
FXGB_CHICK
PR4B_HUMAN
PR4B_MOUSE
HNF6_HUMAN
HNF6_RAT
SELP_RAT
SELP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HUNB DROIX
HUNB DROCY
HUNB DROLI
MOX2 HUMAN
MOX2 MOUSE
MOX2 RAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PC_DROME
NOR2_RAT
ZP12_BRARE
PO33_BRARE
NR43_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                       - protein search, using sw model
                                                                                                                                                                                                                              US-10-018-103A-5
136
1 KHKHHKHHKHKHKHKHK 20
                                                                                                                                                                                                                                                                                                                BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SwissProt_41:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query
Match
                                                                                                                                                                                                                                Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81.5
81.5
81.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         79.5
79.5
79.5
79.5
                                                                                                                                                                                                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score
                                                                                                       OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Database :
                                                                                                                                                                                                                                                                                                                                                                              Searched:
                                                                                                                                                                                                                                                                           Sequence:
                                                                                                                                                Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Result
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                           drosophila
drosophila
schistosoma
drosophila
drosophila
                                                                                                                                                                                                                   drosophila
homo sapien
drosophila
                                                schizosacch
                                                                         drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDLINE=89315821; PubMed=2568636;
MEDLINE=89315821; PubMed=2568636;
MEDLINE=89315821; PubMed=2568636;
The homeotic gene spalt (sal) evolved during Drosophila speciation.";
Proc. Natl. Acad. Sci. U.S.A. 86:5483-5486(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-MAY-1991 (Rel. 18, Created)
01-MAY-1991 (Rel. 18, Last sequence update)
01-MAY-1991 (Rel. 31, Last annotation update)
01-FEB-1996 (Rel. 31, Last annotation update)
Protein spalt-accessory.
SALA OR SAL.
Drosophila simulans (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Endopterygota;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Drosophila melanogaster (Fruit fly).
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
P51179
Q48251
Q46258
Q46258
Q46252
P08016
Q46254
P10035
Q14594
Q26307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 65.4%; Score 89; DB 1; Length 139; Best Local Similarity 64.7%; Pred. No. 0.00027; Matches 11; Conservative 4; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FINASSE; FBGN0012892; Dsim\sala.
SEQUENCE 139 AA; 14145 MW; AED5D51561C229CF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAY-1991 (Rel. 18, Created)
01-MAY-1991 (Rel. 18, Last sequence update)
01-FEB-1995 (Rel. 31, Last annotation update)
Protein spalt-accessory.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       139 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          142 AA
                                                                                                                                                                                                                                                                                                                                                  ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           or send an email to license@isb-sib.ch).
NR43 RAT
HPN HELPY
YBEZ SCHPO
HUNB DROMH
HUNB DROSO
HUNB DROPO
HUNB DROPO
HUNB DRODA
HUNB DRODA
HUNB DRODA
HUNB DROMB
PGCS SCHWA
NA DROMB
NGMA DROMB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           123 EHHEHHGHHEHHGHRH 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  КНИКНИКНИКНИМ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M21227; AAA28877.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
```

Gaps

ö

```
SUGURNER, EXURAN N. A. LEOUTONIN CLARS & C. M. SCRANB G. A. GOCCAYNE J. D. C. SCHARES S. HOLT R.A. EVARRE R.A. GGALER F. P. AMARIAGE S.D. CELLINE S. C. CHARIGE R.A. GENERER S. B. HOLT R.A. BABDUTHER M. HENGERON S.N. AMARIAGES P. G., SCHARES S.E. HILD P. M. LEANE B. M. LEANE S.E. R. AMARIAGES F. G. CHARING M. HENGERON S.N. SULTON G.G., WONTHMEN J. R. HALL G. M. PARING M. M. BERNER Y. H. G. HARLE R. M. HELL G. M. HELL G. M. MALOS G.L.G. MAN M. BERNER M. M. MURDHY B. MURDHY B. MONDER M. M. MENDHY B. MURDHY M. MURD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=96038094; PubMed=7555710; Treisman J.E., Lai Z.-C., Rubin G.M.; Treisman J.E., Lai Z.-C., Rubin G.M.; Treisman J.E., Lai Z.-C., Rubin G.M.; Sopritsighted acts in the decapentaplegic pathway in Drosophila eye development and has homology to a mouse TGF-beta-responsive gene."; Development 121:2835-2845(1995).
                                                                                                                                                                                                                                                                                   28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Bunched protein, class 2/class 3 isoforms (Shortsighted protein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. (ISOFORM CLASS 2), SUBCELLULAR LOCATION, AND
                                                                                                                                                                                                                                                                                                                                                       Bunched protein, class 2/class 3 isoforms (Shortsighted protein BUN OR SHS OR CG5461.

Drosophila melanogaster (Fruit fly).

Bukaryota, Metazoa, Arthropoda; Hexapoda; Insecta; Pterygota; Roptera; Endoptera; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A. (ISOFORMS CLASS 2 AND CLASS 3).
  5
                                                                                                                                                                                                                                     PRT; 1211 AA.
     Mismatches
        4
                                                                                    :||:|| ||:||.||:|
126 EHHEHHGHHEHHGHRH 142
                                                         13
                                                         з кинкинкинкинки
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUE=Eye-antennal disk;
                                                                                                                                                                                                                                             STANDARD;
        Conservative
                                                                                                                                                                                                                                                                        Q245<u>2</u>3; Q9VK78; Q9VK79;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=7227;
          11;
                                                                                                                                                                                                                                               BUN2 DROME
                                                                                                                                                                                                                    BUNZ DROME
                Matches
                                                                                                                                                                                                                                                                        셤
                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=89315821; PubMed=2568636;
REULET D., Schuh R., Jaeckle H.;
Reuler D., Schuh R., Jaeckle H.;
"The homeotic gene spalt (sal) evolved during Drosophila speciation.";
Proc. Natl. Acad. Sci. U.S.A. 86:5483-5486(1989).
                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
MEDLINE=89315821; PubMed=2568636;
Reuter D., Schuh R., Jaeckle H.;
Reuter D., Schuh R., Jaeckle H.;
"The homeotic gene split (sal) evolved during Drosophila speciation.";
"The homeotic gene split (sal) evolved during Drosophila speciation.";
Proc. Natl. Acad. Sci. U.S.A. 86:5483-5486(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                       Frei E., Schuh R., Baumgartner S., Burri M., Noll M., Juergens G., Seifert E., Nauber U., Jaeckle H.; "Molecular characterization of spalt, a homeotic gene required for head and tail development in the Drosophila embryo."; EMBO J. 7:197-204(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota, Metazoa, Arthropoda, Hexapoda, Insecta, Pterygota, Neoptera, Endopterygota, Diptera, Brachycera, Muscomorpha, Ephydroidea, Drosophilae, Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          65.4%; Score 89; DB 1; Length 142; 64.7%; Pred. No. 0.00028;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 89; DB 1; Length 142;
Pred, No. 0.00028;
4; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FlyBase; FBgm0012624; Dore\sala.
SEQUENCE 142 AA; 14498 MW; 829A9A586EFB25F1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FlyBase, FBgn0003313; sala.
SEQUENCE 142 AA; 14554 MW; 7386D6C8C12E5044 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-MAY-1991 (Rel. 18, Created)
01-MAY-1991 (Rel. 18, Last sequence update)
01-FEB-1996 (Rel. 33, Last annotation update)
Protein spalt-accessory.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              142 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     65.4%; Sco...
64.7%; Pred
4; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             :||:|| ||:|| ||:|
126 EHHEHHGHHEHHGHHRH 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               з кинкинкинки 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M21579; AAA28876.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Drosophila orena (Fruit fly)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; X57474; CAA40712.1; -. PIR; S00262; S00262.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 64.7 es 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR; C33910; C33910.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                               EQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=7233;
                             NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SALA_DROOR
ID SALA_DROOR
AC P21748;
```

Query Match Matches

ઠ g

RESULT 3

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
nervous system morphogenesis, eye development and oogenesis. May be required for the transmission of the depp signal and for a morphogenetic movement of the medulla in the brain that reorients the second optic lobe relative to the first. Plays a role in determining proper dorsal cell fates leading to the formation of the dorsal appendages.

SUBCELLULAR LOCATION: Nuclear and cytoplasmic.

ALTERNATUR PRODUCTS:

ALTERNATUR PRODUCTS:

Comment=Experimental confirmation may be lacking for some
                                                                                                                                                                                                                                 Insoid=024523-2; Sequence=VSP 006670; SIMILARITY: BELONGS TO THE TSC-22/DIP/BUN FAMILY. CAUTION: Ref.2 sequence differs from that shown due to erroneous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               protein; Alternative splicing.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 1211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /FTIGAVSP 006670.

K -> E (IN REF. 1).
MISSING (IN REF. 1).
QO.TSAA -> TS (IN REF. 1).
QO.TSAA -> TS (IN REF. 2).

MW; 9925A9159A7051B0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Missing (in isoform Class 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 87; DB 1; Length 121
Pred. No. 0.0035;
2; Mismatches 4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLY-ALA.
POLY-ALA.
LEUCINE-ZIPPER.
POLY-ALA.
                                                                                                                                                                             IsoId=Q24523-1; Sequence=Displayed;
                                                                                                                                                                                         Name=Class 1;
IsoId=Q24522-1; Sequence=External;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ALA-RICH
POLY-SER
POLY-SER
POLY-SER
POLY-HIS
POLY-HIS
POLY-GIN
POLY-RICH
POLY-RICH
POLY-RICH
POLY-RICH
POLY-RICH
POLY-RICH
POLY-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Flyase; FBGN0010460; bun.
InterPro; IPR000580; TSC-22_Dip_Bun.
Pfam; PF01166; TSC-22_1 1.
ProDom; P0007122; TSC-22_Dip_Bun; 1.
PROSITE; PS01289; TSC-22_Dip_Bun; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL, AE003636, AAF53200.1, ALT_SEQ.
EMBL, AE003636, AAF53201.1, ALT_SEQ.
PIR, T13804, T13804.
HSSP, P80220, 1DIP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RUSIIE; FSU1202, 1000, Nuclear Iranscription regulation; Nuclear Nuclear 15 31 POLY-
                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; L42512; AAC41608.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 64.7%;
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              125299
                                                                                                                                                                                                                                                                                gene model prediction.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1014
1090
1201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1211 AA;
                                                                                                                                                                                                                      Name=Class 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1194
                                                                                                                                                isoforms;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
DOMAIN
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseaisb-sib.ch).
                                                                                                                                                                                                                                                                               MEDIJINE-812753136; PubMed=3301325;
MEDIJINE-812753136; PubMed=3301325;
Anders R.F., Stahl H.-D., Crewther P.E., Scanlon D.B., Brown G.V.,
Anders R.F., Kemp D.J.;
"The Complete sequence of the gene for the knob-associated histidine-
rich protein from Plasmodium falciparum.";
FUNCTION: KAHRP MIGHT MIMICK HUMAN HISTIDINE-RICH GLYCOPROTEINS
TO ANCHOR HOST THROMBOSPONDIN OR A PARASITE ANALOG IN A BINDING
COMPLEX WITH THE ENDOTHELIAL CELL RECEPTOR.
-- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF THE MEMBRANE OF INFECTED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL.
KNOB-ASSOCIATED HISTIDINE-RICH PROTEIN.
HIS-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mus musculus (Mouse)..
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC. . .) (POTENTIAL)
F69E1B9CCF98F946 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                  Planmodium falciparum (isolate NF7 / Ghana).
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 85, DB 1; Length 657;
Pred. No. 0.0033;
3; Mismatches 4; Indels
                                                                                                                                                        01-JAN-1988 (Rel. 06, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
Knob-associated histidine-rich protein precursor (КАНRP)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-NUV-1997 (Rel. 35, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
Forkhead box protein B2 (Transcription factor FKH-4).
FOXE2 OR FKH4.
                                                                                                        657 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  428 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     42 N.
71941 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (Rel. 35, Created)
(Rel. 35, Last seqn
(Rel. 39, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   з кникникникнких 20
              306 HHHHHHHHHHHQHHNHQ 322
                                                                                                                                         06, Created)
06, Last seq
никинкинкникик 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; Y00060; CAA68268.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Malaria, Repeat, Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34
657
123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      657 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR; A29454; A29454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
                                                                                                                                       01-JAN-1988 (Rel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ERYTHROCYTES.
                                                                                                                                                                                                                                                   NCBI_TaxID=5842;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1997
                                                                                                         KNOB PLAFN
P06719;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOUSE
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FXB2_MOUSE
ID FXB2_MOU
AC Q64733;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                           PLAFN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REXXOCOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ઠે
ઠે
```

ö

Gaps

```
MEDLINE=87270765; PubMed=3038111;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                               Malaria.
SEQUENCE
                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
REPEAT
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROPEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstationthe European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
              Kaestner K.H., Schuetz G., Monaghan A.P.; "Expression of the winged helix genes fkh-4 and fkh-5 defines domains in the central nervous system.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                      MEDLINE-93361500; PubMed-7689224;
Kaestner K.H., Lee K.H., Schloendorff J., Hiemisch H.,
Monaghan A.P., Schuetz G.;
"Six members of the mouse forkhead gene family are developmentally
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-MAY-1991 (Rel. 18, Last annotation update)
Histidine-rich protein.
Plasmodium falciparum (isolate fcm17 / Senegal).
Bukaryota, Alveolata, Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5845;
                                                                                                                                                                                                                                                                                                                                                                             HSSP; 02345; D24742;
HSSP; 02345; D2442;
TRANSFAC; T02442;
MGD; MGI:1347468; Foxb2.
MGD; MGI:1347468; Foxb2.
InterPro; PR001766; TF Fork head.
Ffam; PR00256; Fork head; 1.
PRINTS; PR0053; FORKHBAD.
ProDom; PD000425; TF Fork head; 1.
PROSITE; PS00657; FORK HEAD 1; 1.
PROSITE; PS00657; FORK HEAD 1; 1.
PROSITE; PS00658; FORK HEAD 2; 1.
PROSITE; PS0093; FORK HEAD 2; 1.
PROSITE; PS0093; FORK HEAD 2; 1.
PROSITE; PS0093; FORK HEAD 3; 1.
PROSITE; PS0093; FORK HEAD 3; 1.
PROSITE; PS0093; FORK-HEAD 3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61.0%; Score 83; DB 1; Length 428; 68.8%; Pred. No. 0.0036; tive 0; Mismatches 5; Indels
                                                                                                                                                                     Proc. Natl. Acad. Sci. U.S.A. 90:7628-7631(1993).
-!- SUBCELLULAR LOCATION: Nuclear.
-!- DEVELOPMENTAL STAGE: EXPRESSED DURING EMBRYOGENESIS.
-!- SIMILARITY: Contains 1 fork-head domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB8A8EFD1E94AB10 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (Rel. 13, Created)
(Rel. 13, Last sequence update)
(Rel. 18, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                82 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLY-ALA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLY-ALA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLY-ALA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLY-ALA
         MEDLINE=97014266; PubMed=8861101;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              45170 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | || || || || || || 11 || 11 || 137 НРНИРНЯАНИННИН 152
                                                                                                                                                                                                                                                                                                                                           EMBL; X92591; CAA63335.1; -. EMBL; X71942; CAA50744.1; -. PIR; D47746; D47746.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         нкинкинкинкин 17
                                                          Mech. Dev. 55:221-230(1996)
                                                                                   SEQUENCE OF 4-114 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                162
172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             428 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-JAN-1990
01-JAN-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PLAFS
                                                                                                                                                               regulated.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P14586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HRP3 PLAFS
TID HRP3 P
AC 01-4586
DT 01-478N
DT 01-78N
DT 01-MAX
DE H18tid
OS P18sm
OX NCBL T
RN (1)
RP SEQUEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 7
          용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ठ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -i- MISCELLANBOUS: IN THE INTRAERYTHROCYTIC STAGES OF DEVELOPMENT OF P.LOPHURAE IN DUCKS, THERE IS.A SYNTHESIS OF A MAJOR PROTEIN THAT ACCUMULATES TO COMPRISE AT LEAST 50% OF THE CELLULAR MASS: THE HISTIDINE RICH PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
Ravetch J.V., Feder R., Pavlovec A., Blobel G.; "Primary structure and genomic organization of the histidine-rich protein of the malaria parasite Plasmodium lophurae."; Nature 312:616-620 (1984).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HISTIDINE-RICH GLYCOPROTEIN.
N-LINKED (GLCNAC. . .) (PROBABLE).
Z X 16 AA TANDEM REPEATS.
16-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Plasmodium lophurae.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5853;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 X 17 AA TANDEM REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82 AA; 10578 MW; 7CE2EA69F2FC1E8C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 82.5; DB 1;
Pred. No. 0.00087;
4; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       351 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-AUG-1987 (Rel. 05, Created)
13-AUG-1987 (Rel. 05, Last sequence upo
15-JUL-1999 (Rel. 38, Last annotation u
Histidine-rich glycoprotein precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR; A22692; KGZQHL.
Malaria; Repeat; Glycoprotein; Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDLINE=85061618; PubMed=6095114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2 НИНИИНИНИНИНИНКИК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |:|| ||:|| || ||:||:
22 HRHH-HHRHHHRHQ 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; X01469; CAA25698.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; M17028; AAA29619.1; -. PIR; A29653; A29653.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             60.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     57.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 8
HRPX_PLALO
ID P04929;
AUG-1987
```

36 HSHNSHHPHRHHHHH 54

용

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EWBL outstation the European Bioinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a alicense agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1;
                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                     Avian sarcoma virus (strain 31) (ASV31).
Viruses; Retroid viruses; Retroviridae; Avian type C retroviruses.
VCBI_TaxID=35270;
                                                                                                                                                                                                                                                                                                                                                                                                                                             Li J., Vogt P.K.;
"The retroviral oncogene qin belongs to the transcription factor family that includes the homeotic gene fork head.";
Proc. Natl. Acad. Sci. U.S.A. 90:4490-4494(1993).
-!- SUBCELLUIAR LOCATION: UNclear.
-!- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-QIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  <del>.</del>
                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1; Length 387;
                                                                                                            Length 351;
                                                                                                          60.3%; Score 82; DB 1; Length 351
68.8%; Pred. No. 0.0039;
ive 0; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6; Indels
17-1.
17-2.
2 X 15 AA TANDEM REPEATS.
15-1.
15-2.
                                                                  18 X 10 AA TANDEM REPEATS
D19A48D47D890453 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEA902F50FFE42F9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 81.5; DB 1;
Pred. No. 0.0049;
                                                                                                                                                                                                                                                                                          15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Fransforming protein Qin (Oncogene Qin).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -!- SIMILARITY: Contains 1 fork-head domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TRANSFAC; T0182; ...
InterPro. IPR001766; TF Fork_head.
Pfam, PF00250; Fork_head; 1.
PRINTS; PR00053; FORKHEAD.
PRODOM; PD000445; TF Fork_head; 1.
SWART; SW00339; FH; T.
PROSITE; PS00657; FORK_HEAD_1; 1.
PROSITE; PS00658; FORK_HEAD_2; 1.
PROSITE; PS00659; FORK_HEAD_2; 1.
PROSITE; PS00659; FORK_HEAD_3; 1.
DNA-binding; Nuclear protein; Oncogene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLY-HIS.
POLY-PRO.
POLY-ALA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FORK-HEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLY-ALA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; L10719; -; NOT_ANNOTATED_CDS
HSSP; Q63245; 2HFH.
                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
MEDLINE=93281605; Pubmed=8099441;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           42283 MW;
                                                                                  44032 MW;
                                                                                                                                                                                           215 нинининискини 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    59.9%;
                                                                                                                                                                 4 ННЮНИМНИМНИМ 19
                                                                                                                       Local Similarity
nes 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches 12; Conservative
                                                                                                                                                                                                                                                                STANDARD;
 107
123
153
138
153
351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           387 AA;
                                                                                351 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLYPROTEIN.
                                                                                                                                                                                                                                                               QIN AVIS3
P56<u>2</u>60;
                                                                   DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local
 REPEAT
REPEAT
DOMAIN
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                  V-OIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                  QIN_AVIS3
                                                                                                                                       Matches
                                                                                                                                                                                                                                      RESULT 9
                                                                                                                                                                                                                                                                             222223
                                                                                                                                                                  8
                                                                                                                                                                                         g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.lsb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-White leghth.

STRAIN-White leghth.

STRAIN-White leghth.

STRAIN-White leghth.

STRAIN-White leghth.

A Yuasa J., Hirano S., Yamagata M., Noda M.;

Yuasa J., Hirano S., Yamagata M., Noda M.;

T "Visual projection map specified by topographic expression of

T "Visual projection map specified by topographic expression of

T "Visual projection map specified by topographic expression of

THEN TARGET GAS-655 (1996).

C THEN TARGET GENES.

THEIR TARGET GAN BE DETECTED IN REGIONS INCLUDING

PRINORDIAL RETINA AND NEUROEPETHELIUM BY EMBRYONIC DAY 2 (E2). AT

E3, EXPRESSED IN THE NAGAL RETINA AND PIGMENT EPITHELIUM AS WELL

C GANGLION CELLS. LEVELS BEGIN TO DECLINE FROM E4 AND ALMOST

C DISAPPERR BY STRAIN.

C DISAPPERR BY STRAIN.

C DISAPPERR BY STRAIN.

C DISAPPERR BY STRAIN.
                                                                                                                                                                                                                                                                                          Gallus gallus (Chicken).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                       15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Forkhead box protein G1B (Forkhead-related protein FKHL1)
Forshead ON G1N (10-62-5) (CEQ 3-1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; 136814; AAA66954.1; -...
EMBL; U47275; AAB08466.1; -...
PIR, AAS909; ASS909.
HSSP; Q63245; ZHPH.
TRANSFAC; T01833; -...
InterPro; PR001766; FORK head; 1...
PRINTS; PR00053; FORKHEAD.
PRODOM; PR000425; FT FORK head; 1...
PROSITE; PS000659; FORK HEAD.1; 1...
PROSITE; PS000659; FORK HEAD.1; 1...
PROSITE; PS00059; FORK HEAD.2; 1...
PROSITE; PS00099; FORK HEAD.3; 1...
TRANSCTIPLION SAMEN SECONDAL SECONDA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDLINE=95132616; Pubmed=7831308;
Chang H.W., Li J., Kretzschmar D., Vogt P.K.;
Avian cellular homolog of the qin oncogene.";
Proc. Natl. Acad. Sci. U.S.A. 92:447-451(1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY: Contains 1 fork-head domain.
451 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          protein; Proto-oncogene.
15 54 POLY-HIS.
55 58 POLY-PRO.
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=9031;
CHICK
                                                                                                                                                                                                                                                                                                                                                                                                     Gallus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
```

```
Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arawai J., Shingawa A., Shibata K., Konno H., Adachi J., Fukuda S., Arawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Gojobori T., Bono H., Rasukawa T., Saito R.,
                                                                                                                                                                                                                                                                                                                                                                                  NP BIND
BINDING
ACT SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
PR4B_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                              ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --- FUNCTION: Has a role in pre-mRNA splicing. Phosphorylates SF2/ASF.
--- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
--- SUBCELLIVIAR LOCATION: Nuclear.
--- SUBCELLIVIAR LOCATION: Nuclear.
--- TISSUE SPECIFICITY: Ubiquitous.
--- PTM: Phosphorylated by Clk1.
--- PTM: Phosphorylated by Clk1.
--- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
                                                                                                                                                                                                                          013523; QBTDP2; Q9GQT7; Q9UEE6; 15-JUL-1998 (Rel. 36, Created) 28-FEB-2003 (Rel. 41, Last sequence update) 28-FEB-2003 (Rel. 41, Last annotation update) serine/threonine-protein Kinase PRP4 homolog (EC 2.7.1.37) (PRP4 pre-PRP4B OR PRP4H OR PRP4K OR KIAA0536.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Functional analysis of the fission yeast Prp4 protein kinase
involved in pre-mRNA splicing and isolation of a putative mammalian
                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O., "Prediction of the coding sequences of unidentified human genes. IX The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gross T., Lutzelberger M., Wiegmann H., Klingenhoff A., Shenoy S., Kaeufer N.P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ğ
                                                                                                                                                                                                                                                                                                                                Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                              J. Biol. Chem. 276:32247-32256(2001).

[2]

SEQUENCE FROM N.A.

MEDLINE=22072586; PubMed=12077342;

Dellaire G., Makarov E.M., Cowger J.J.M., Longman D.,

Sutherland H.G.E., Luhrmann R., Torchia J., Bickmore W.A.;

"Mammalian PRP4 kinase copurifies and interacts with components both the US snRNP and the N-CoR deacetylase complexes.";

Mol. Cell. Biol. 22:5141-5156(2002).
                                                                                               ï
                                                                   Score 81.5; DB 1; Length 451;
Pred. No. 0.0056;
                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                     SECUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE=21402934; PubMed=11418664;
Kojima T., Zama T., Wadda K., Onogi H., Hagiwara M.;
"Cloning of human PRP4 reveals interaction with CIK1.";
J. Biol. Chem. 276:32247-32256(2001).
                               FORK-HEAD.
E9E5B407D2321B50 CRC64;
                                                                                                                                                                                                                     PRT; 1007 AA.
                                                                                                  Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nucleic Acids Res. 25:1028-1035(1997).
        POLY-ALA.
POLY-ALA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 512-1007 FROM N.A.
MEDLINE=97250302; PubMed=9102632;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TISSUE=Brain;
MEDLINE=98290545; PubMed=9628581;
                                                                                                  ·
6
                                                                                                                           2 НХНН-КНИКНИКНИКН 19
                                                                                                                                              36 HSHHNSHHPHHHHHHHH 54
                                                48856 MW;
                                                                       59.98;
                                                                                    63.2%;
                                                                  Query Match
Best Local Similarity 63.29
watches 12; Conservative
                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                               Homo sapiens (Human)
         64
103
142
451 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEOUENCE FROM N.A.
                                                                                                                                                                                                                       PR4B HUMAN
                       DOMAIN
DNA BIND
SEQUENCE
                                                                                                                                                                                              RESULT 11
PR4B_HUMAN
                                                                                                                                                                                                                                     셤
            5558
                                                                                                                               ð
```

```
÷.
loved. Usage by and for commercial (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                GK; Q13523; -.

R MIM; 602336; -.

R MIM; 602336; -.

R GV; Q13523; -.

R GV; Q0338; -.

R GV; C0306468; Piprotein amino acid phosphorylation; TAS.

GO; C006468; Piprotein amino acid phosphorylation; TAS.

R InterPro; IPR001245; Tyr_pkinase.

R InterPro; IPR001245; Tyr_pkinase.

R Probom; PD00009; Pkinase; 1.

R PRART; SM00220; STKC; 1.

R SMART; SM00220; STKC; 1.

R PROSITE; PS00109; PROTEIN KINASE_ATP; FALSE_NEG.

R PROSITE; PS00109; PROTEIN KINASE_DOM; 1.

R PROSITE; PS00109; PROTEIN KINASE_DOM; 1.

R PROSITE; PS001019; PROTEIN KINASE_DOM; 1.

R NEAL STANCESBALDS; MRNA splīcīng; Transferase;

M MRNA procesbalng; MRNA splīcīng; Transferase;

M SCTING/Lhreonine-protein kinase; ATP-binding; Phosphorylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 1007;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                45ECF73ABB56D17C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATP (BY SIMILARITY).
ATP (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 81.5; DB 1;
Pred. No. 0.012;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 КНКН---НКННКН------НКН-НКНК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ARG/LYS-RICH.
PROTEIN KINASE.
    modified and this statement is not removed. entities requires a license agreement (See h or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BY SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HIS-RICH
                                                                                                                 EMBL; AF203465; AAM19101.1; --
EMBL; AB011109; BAAZ5462.1; ALT_INIT.
EMBL; U48795; AAB03268.1; --
HSSP; P24941; 1AQ1.
Genew; HGNC:17346; PRPF4B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         116972 MW;
                                                                                              EMBL; AY029347; AAK38155.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   59.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 54.5
Matches 18; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         78
496
1006
701
717
815
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            39 49
687 100
693 70
717 71
815 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nuclear protein.
DOMAIN 40
```

PR4B MOUSE STANDARD; PRT; 1007 AA.

Q61136; O88378; QBR4Y5; Q9CTL9; Q9CTT0;

15-JUL-1998 (Rel. 36, Created)

28-FEB-2003 (Rel. 41, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)

Serine/threonine-protein kinase pRel homolog (BC 2.7.1.37) (PRP4 premRNA processing factor 4 homolog) (Pre-mRNA processing factor 4 homolog) (Pre-mRNA protein kinase).

MUS MUSCULUS (MOUSE). Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus Dellaire G., Makarov E.M., Cowger, J.J.M., Longman D., Sutherland H.G.E., Lubrmann R., Torchia J., Bickmore W.A.; wammalian PRP4 kinaen copurifies and interacts with components both the US snRNP and the N-CoR deacetylase complexes."; STEAIN-C57BL/6J; IISSUE-Hippocampus, and Urinary bladder; MEDLINE-21085660; PubMed-11217851; MEDLINE=22072586; PubMed=12077342; SEQUENCE OF 1-279 FROM N.A. SEQUENCE FROM N.A. NCBI TaxID=10090;

```
55558
                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                              ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Marsuo Y., Nikaido I., Peoleole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.F., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Saboki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wanshima W., Kawaji H., Kohtsuki S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Functional analysis of the fission yeast Prp4 protein kinase
involved in pre-mRNA splicing and isolation of a putative mammalian
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 185-1007 FROM N.A.
MEDLINE=98369054; PubMed=9701556;
Tate P., Lee M., Tweedie S., Skarnes W., Bickmore W.;
"Capturing novel mouse genes encoding chromosomal and other nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-97250302; PubMed=9102632;
Gross T., Lutzelberger M., Wiegmann H., Klingenhoff A., Shenoy S.,
Kaeufer N.P.;
                                                                                                                                                                                                                                                                                                                                                                                                    "Functional annotation of a full-length mouse cDNA collection.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Proposition Processing Processing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ARG/LYS-RICH,
PROTEIN KINASE.
ATP (BY SIMILARITY).
ATP (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HIS-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MGD; MGI:109584; Prpf4b.
GO; GO:0005694; C:chromosome; IDA.
InterPro; IPR002019; Prot kinase.
InterPro; IPR002290; Ser_Chr_Dkinase.
InterPro; IPR001245; Tyr_Dkinase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Cell Sci. 111:2575-2585(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AF283466; AAM19102.1; -.
EMBL; AK021579; BAB32337.1; -.
EMBL; AK021274; BAB32358.1; -.
EMBL; AF033663; AAG32042.1; -.
EMBL; U48731; AAB03269.1; -.
HSSP; P24941; 1AQ1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 512-1007 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                    Nature 409:685-690(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             497
1006
701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                             Hayashizaki Y.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NP BIND
BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                ä,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=99420592; PubMed=10491763; Moeller A.M., Ek J., Durviaux S.M., Urhammer S.A., Clausen J.O., Moeller A.M., Ek J., Durviaux S.M., Urhammer S.P., Podersen J.O., Eiberg H., Hansen T., Rousseau G.G., Lemaigre F.P., Podersen O.; Hepatcoyte nuclear factor-6: associations between genetic variability and type II diabetes and between genetic variability and estimates of insulin secretion.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Furuta H., Horikawa Y., Wang Y.-Q., Bell G.I.; "Isolation and characterization of the human hepatocyte nuclear factor
                                                                                                                                                                                                                                                                                         Q9UBĞO; Q99744; Q9UMR6;
16-OCT-2001 [Rel. 40, Created)
16-OCT-2001 [Rel. 40, Last sequence update)
28-FEB-2003 [Rel. 41, Last annotation update)
Hepatcoyte nuclear factor 6 (HNF-6) (One cut domain family member 1).
ONECUTI OR HNF6A OR HNF6.
                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
NCBI_TaxID=9606,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Samadani U., Costa R.H.; "Yeast one-hybid cloning of the partial human cDNA for hepatocyte
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. Full G.I.; Furuta H., Wang Y.-Q., Bell G.I.; "The sequence of human mRNA for the hepatocyte nuclear factor-6
                                                                                                                                13;
                                                                                                 DB 1; Length 1007;
                                                                                                                                Indels
                                                                 MW; 18E6C3C43BE7AB4C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6 gene.";
Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
BY SIMILARITY.

SKS -> IPG (IN REF. 3).

K -> I (IN REF. 3).

P -> L (IN REF. 4).
                                                                                                    0.012;
2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1-368 FROM N.A., AND VARIANT ALA-75.
                                                                                                                                                                                    48 KHKHRSKHKKHKHSSEEDRDKKHKHKHKKHK 80
                                                                                                                                                                1 KHKH---HKHHKH------HKH-HKHHKHK 20
                                                                                                                                0; Mismatches
                                                                                                 Score 81.5;
                                                                                                                 Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Diabetologia 42:1011-1016(1999).
815 815
185 187
223 223
633 633
1007 AA; 116947 M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 174-465 FROM N.A.
                                                                                               59.9%;
                                                                                                                                18; Conservative
                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens (Human)
                                                                                                              Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                             HUMAN
ACT SITE
CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                 Query Match
                                                                                                                                Matches
```

```
369
444
72
138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               465 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HNP6_RAT
AC P10512; OBD
DT 16-OCT-200
DT 16-OCT-200
DT 28-FBB-200
DE HepatoCyte
GN ONECUTI OF
GN ONECUTI OF
CO MAMMAlia;
OC MAMMAlia;
O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Activator.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DNA_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઢ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rausa F., Samadani U., Ye H., Lim L., Fletcher C.F., Jenkins N.A., Copeland N.G., Costa R.H.; Lim L., Fletcher C.F., Jenkins N.A., The cut-homeodomain transcriptional activator HNF-6 is coexpressed with its target gene HNF-3 beta in the developing murine liver and pancreas.";

Pancreas.";

Dev. Biol. 192:228-246(1997).

-!- FWUNTION: TRANSCRIPTIONAL ACTIVATOR. BINDS THE CONSENSUS SEQUENCE S'-DHWATTGAYTWMD-3' ON A VARIETY OF GENE PROMOTERS SUCH AS THOSE OF HNF3B AND TTR. IMPORTANT FOR LIVER GENES TRANSCRIPTION (BY SIMILARITY).
R EMBL; AF035581; AAD02033.1; -.

R EMBL; AF035580; AAD02033.1; -.

R EMBL; U964173; AAD02836.1; -.

R REML; V17739; CAB50765.1; -.

R RABAPC; T03286; -.

R GORDW; HGWC:8138; ONECUT1.

MIM; 604164; -.

MIM; 604164; -.

R GO; GO:0006354; C:nucleus; NAS.

R GO; GO:0015563; F:transcriptional activator activity; NAS.

R GO; GO:0016563; F:transcriptional activator activity; NAS.

R InterPro; IPR007108; Cut_homeo.

R InterPro; IPR003350; Homeo. CUT.

R InterPro; IPR003356; Homeo. CUT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HNF6 MOUSE STANDARD; PRT; 465 AA.
008755;
16-OCT-2001 (Rel. 40, Last sequence update)
28-FSB-2003 (Rel. 41, Last annotation update)
48-FSB-2003 (Rel. 41, Last annotation update)
48-FSB-2003 (Rel. 41, Last annotation update)
48-FSB-2003 (Rel. 41, Last annotation update)
60-FSB-2003 (Rel. 41, Last annotation update)
61-FSB-2003 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus
                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001356; Homeobox.
Pfam; PF02376; CUT; 1.
Pfam; PF00046; homeobox; 1.
Probom; PD000010; Homeobox; 1.
SMART; SM00389; HOX; 1.
SMART; SM00389; HOX; 1.
PROSITE; PS00027; HOMEOBOX 1; FALSE_NEG.
PROSITE; PS50071; HOMEOBOX 2; 1.
Transcription regulation; Homeobox; DNA-binding; Nuclear protein; Activator; Polymorphism.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ÷
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                58.8%; Score 80; DB 1; Length 465; 55.0%; Pred. No. 0.0085; arive 1; Mismatches 8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HOMEOBOX.
POLY-HIS.
POLY-HIS.
POLY-ER.

P -> A.

FITIG=VAR. 010729.
A -> T (IN REF. 3 AND 4).
C -> H (IN REF. 4).
R -> K (IN REF. 4).
K -> Q (IN REF. 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F47F78957A6ECFC2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARIII).
SUBUNIT: BINDS DNA AS A MONOMER (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                465 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121 КРРНЫННИННИННИЯ 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 КНКНИКНИКНИКНКИК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         51023 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 55.0
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            465 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE=Liver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DNA_BIND
DNA_BIND
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   P70512; 088755; 10.0 Created)
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
18-Past-2003 (Rel. 41, Last annotation update)
18-Past-2003 (Rel. 41, Last annotation update)
18-COCT-2001 (Rel. 41, Last annotation update)
18-COCT-2001 (Rel. 40, Last factor 6 (HNF-6) (One cut domain family member 1).
18-COCT-2001 (Rel. 40, Last factor 6 (HNF-6) (One cut domain family member 1).
18-COCT-2001 (Rel. 40, Last annotation family member 1).
19-COCT-2001 (Rel. 40, Last annotation family member 1).
10-COCT-2001 (Rel. 40, Last annotation family member 1).
10-COCT-2001 (Rel. 40, Last annotation family member 1).
11-COCT-2001 (Rel. 40, Last annotation family member 1).

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Hepatocyte nuclear factor 6, a transcription factor that contains novel type of homeodomain and a single cut domain."; proc. Natl. Acad. Sci. U.S.A. 93:9460-9464(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. (ISOFORM ALPHA).
STRAIN=Sprague-Dawley; TISSUE=Liver;
MEDLINE=96382488; PubMed=8790352;
Lemaigre F.P., Durviaux S.M., Truong O., Lannoy V.J., Heuan J.J.,
Rousseau G.G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROSITE; PS00027; HOMEOBOX 1; FALSE_NEG.
PROSITE; PS50071; HOMEOBOX 2; 1.
Transcription regulation; Homeobox; DNA-binding; Nuclear protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN=Wistar;
MEDLINE=98256275; PubMed=9593691;
Lannoy V.J., Buerglin T.R., Rousseau G.G., Lemaigre F.P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  58.8%; Score 80; DB 1; Length 465
55.0%; Pred. No. 0.0085;
tive 1; Mismatches 8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSFAC; T03296;

R WGD; WGT:1196423;

RGO; GO:0003154; P:cell differentiation; IMP.

GO; GO:0003154; P:cell differentiation; IMP.

GO; GO:0007432; P:endoderm development; IMP.

R GO; GO:00007432; P:endoderm development; IMP.

R O; GO:0000760; P:glucose metabolism; IMP.

R InterPro; IPR001369; Homeoo_CUT.

R InterPro; IPR001356; Homeobox.

R Pfam; PF00376; CUT; 1.

R Pfam; PF00046; Homeobox; 1.

R PCDOM; PR000010; Homeobox; 1.

R PAGN; SM00389; HOX; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               460 POLY-SER.
50952 MW; A770D27DD5AAC896 CRC64;
-1- SUBCELLULAR LOCATION: Nuclear.
-1- SIMILARITY: Contains 1 CUT domain.
-1- SIMILARITY: BELONGS TO THE CUT HOMBOBOX FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HOMEOBOX.
POLY-HIS.
POLY-HIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. (ISOFORM BETA).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 КНКНИКНИКНКИКНКИК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; U95945; AAB53863.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 55.0 es 11; Conservative
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                ISOId=P70512-2; Sequence=VSP 002312;
-1- TISSUE SPECIFICITY: EXPRESSED IN LIVER, BRAIN, SPLEEN AND TESTIS.
-1- SIMILARITY: Contains 1 CUT domain.
-1- SIMILARITY: BELONGS TO THE CUT HOMEOBOX FAMILY.
                                                                                                      Samadani U., Costa R.H.; "The transcriptional activator hepatocyte nuclear factor 6 regulates
"Isoforms of hepatocyte nuclear factor-6 differ in DNA-binding properties, contain a bifunctional homeodomain, and define the new ONECUT class of homeodomain proteins.";
J. Biol. Chem. 273:13552-13562[1998].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A -> AESAMGGSVPSLRITSGGPQLSVPPLP (in isoform Beta).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ProDom; PD000010; Homeobox; 1.
SMRAT; SW00389; HOX; 1.
PROSITE; PS00027; HOXEOBOX 1; FALSE_NEG.
PROSITE; PS00071; HOWEOBOX 2; 1.
PROSITE; PS50071; HOWEOBOX 2; 1.
Aramscription regulation; Homeobox; DNA-binding; Nuclear protein; Activative splicing.
DNA_BIND 283 369 CUT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 80; DB 1; Length 465;
Pred. No. 0.0085;
1; Mismatches 8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /FTId=VSP 002312.
465 AA; 51067 MW; BD651267FD7AC896 CRC64;
                                                                                                                                                                                                                                                                                    Event=Alternative splicing; Named isoforms=2;
                                                                                                                                                                                                                                                                                                                 IsoId=P70512-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HOMEOBOX.
POLY-HIS.
POLY-HIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLY-SER
                                                                                      MEDLINE=97042457; PubMed=8887657;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 interPro; INTERO 108; CUL homeo.
InterPro; IPR003350; Hmoeo CUT.
InterPro; IPR001356; HomeoDox.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       58.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; X96553; CAA65389.1; -.
EMBL; X14933; CAA75150.1; -.
TRANSFAC; T03257; -.
TRANSPAC; T03258; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF02376; CUT; 1.
Pfam; PF00046; homeobox; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 55.0
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                     Name=Alpha;
                                                                                                                                                                                                                                                                                                                               Name=Beta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Activator;
DNA BIND
DNA BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
```

Gaps

ö

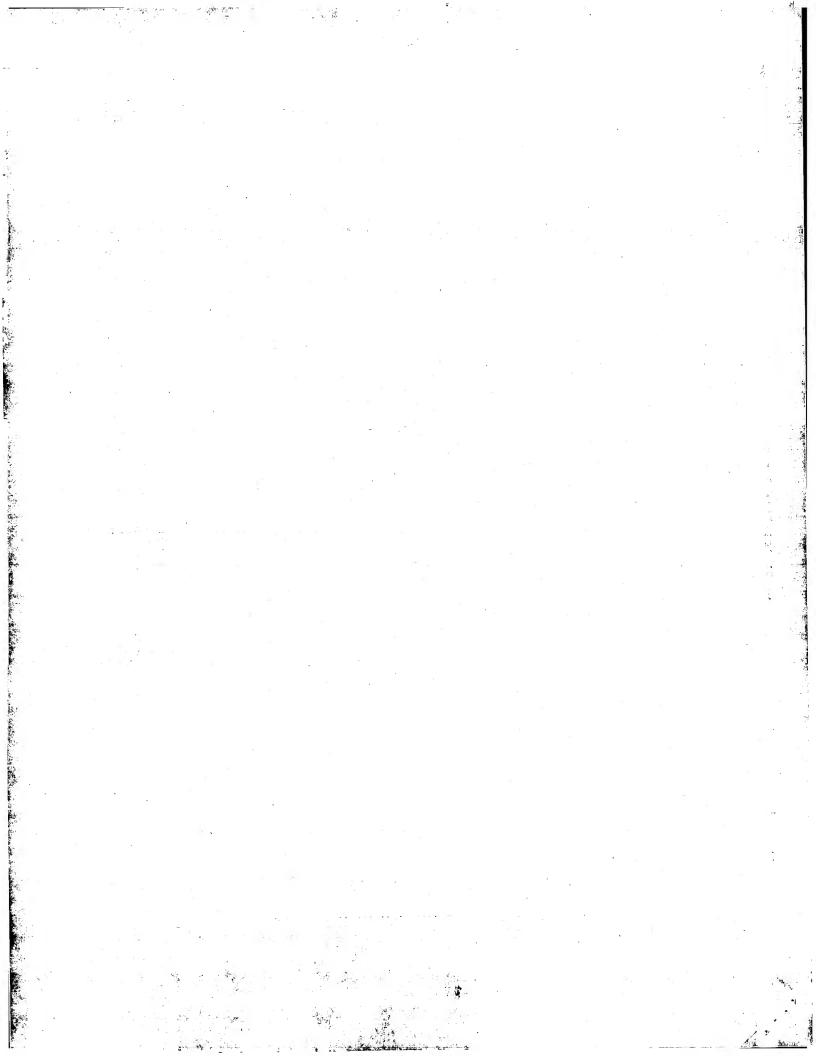
1; Mismatches

Search completed: January 20, 2004, 18:23:56 Job time : 8.66667 secs

121 KPPHRHHHHHHHHHRR 140

1 ХИКИНКИНКИНКИКИК 20

ઠે g



```
January 20, 2004, 18:20:39; Search time 24.6377 Seconds (without alignments) 209.478 Million cell updates/sec
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                           OM protein - protein search, using sw model
                                                                                                                                                                                      US-10-018-103A-5
136
1 KHKHHKHHKHKHKHK 20
                                                                                                                                                                                                                                                                    BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                      Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                      Scoring table:
                                                                                                               Run on:
```

Total number of hits satisfying chosen parameters:

830525 seqs, 258052604 residues

Searched:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries sp_rodent:*
sp_virus:*
sp_vertebrate:*
sp_unclassified:* sp_human:*
sp_invertebrate:*
sp_mammal:*
sp_mhc:* sp_organelle:*
sp_phage:*
sp_plant:* 1: sp_archea:* 2: sp_bacteria:* 3: sp_fungi:* SPTREMBL_23:* 55... 1111... 1112... 114... 115... •• Database

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution. sp_rvirus:*
sp_bacteriap:*
sp_archeap:*

SUMMARIES

Q9v3w5 drosophila
Q8i142 plasmodium
Q9vjj8 drosophila
Q9vjj9 drosophila
Q9vj9 drosophila
Q91447 pseudomonas
Q91447 pseudomonas
Q98474 petromyzon
Q98987 arabidopsis
Q8t2u5 drosophila
Q75c1 drosophila
Q79w5c0 drosophila
Q9w5c0 drosophila
Q9w5c0 drosophila
Q9w5c0 drosophila
Q9ycd0 drosophila
Q9ycd1 drosophila
Q9yck1 drosophila Description 09903W5 0981142 0990739 0910739 0910747 0910729 0917261 077261 077261 099620 099620 099609 Query Match Length DB Score 888 899 901 809 809 809 Result No.

10

Q90zq7 oryzias lat Q9bsn0 homo sapien Q8c4j8 mus musculu	Q969f2 homo sapien Q9d7u9 mus musculu	mus	Q8ve28 mus musculu	Q9i814 brachydanio	Q8mp90 dictyoateli	Q90840 gallus gall	Q9eqn4 mus musculu		Q9ve69 drosophila		Q9vj79 drosophila	Q8mgw0 drosophila	Q95ph4 dictyosteli	Q9bsx0 homo sapien	Q9nz81 homo sapien	Q9ntq8 homo sapien	Q9vbh3 drosophila	Q8mrl5 drosophila	Q8ilv8 plasmodium	Ln.	Q9w493 drosophila	O9spj7 gossypium h		Q8c145 mus musculu
13 Q90ZQ7 4 Q9BSN0 11 Q8C4J8	4 Q969F2	11 Q91Y45	11 Q8VE28	13 Q918L4	S QBMP90	13 Q90840	11 Q9EQN4	5 Q917K2	5 Q9VE69	5 Q95РН8	S Q9VJ79	S Q8MQW0	S Q95PH4	4 Q9BSX0	4 Q9NZ81	4 Q9NTQ8	5 Q9VBH3	S Q8MRLS	S QBILV8	10 Q9XEK5	o		10 023038	11 Q8C145
145 325 449	451	461	461	1281	83	1040	1057	503	652	1147	1365	1544	1318	70	148	148	898	769	954	1033	1059	314	391	765
65.4 65.4 65.4	65.4	65.4	65.4	65.4	64.0	64.0	64.0	63.6	63.6	63.2	63.2	63.2	62.5	61.8	61.8	61.8	61.8	61.4	61.4	61.4	61.4	61.0	61.0	61.0
6 8 8	68 8	6 8	6 6 6	0 80 0 00	87	. 87	87	86.5	86.5	98	98	98	82	84	84	84	84	83.5	83.5	83.5	83.5	83	83	83
17 18 19	20	22	23	52	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

			2				D., R.P.,	, N.S n	or B	8 G.L.G.,	dwin D.,	S.,	o, Jandra I.,		S.M., Dunn P.,	chmann W.,		<i>:</i>	chum K.A.,
	a		Pterygota :omorpha;				Gocayne J. A., Galle	, Hendersc	Chen L.X.	R., Miklo	och C., Bal	Bolshakov	Brottier F Her A., Ch	Davies P.	I., Dietz Junkov B.C.	S., Fleis	Harris M.,	ו, Houck ל Therman	J.A., Ket
186 AA.	sequence update) annotation update)		Eukaryota, Metazoa, Arthropoda, Hexapoda, Insecta, Pterygota, Neoptera, Endopterygota, Diptera, Brachycera, Muscomorpha,				Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.P.,	shburner M	, Zhang O.,	., Nelson (Abril J.F., Agbayani A., An HJ., Andrews-Pfannkoch C., Baldwin D., Ballow B M. Basu A. Bakalev E M.	Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P., Burtis K.C., Bugam D.A., Butler H., Cadieu E., Center A., Chandra I.,	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Duqan-Rocha S., Dunkov B.C., Dunn	K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann	Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,	N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,	M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
PRT; 18	rast sequen Last annota	fly).	ı, Hexapod era, Brach	rosophila.		1132;	Li P.W.	irds S., A	ndell M.D.	J., Helt G	-J., Andr	in B.P., B	ok J., Bro	C., Dave	J Z., Mays	Ferraz C	H., Gu Z.	1 T.J., He	3.H., Ke Z t C., Kra
13,	23,	er (Fruit	rthropoda ta; Dipte	lidae; Dr		Med=10731	S.E., Hol	E., Riche	J.R., Yar	axter E.C	A., An H.	V., Berma	A. Butle	., Dahlke	Downes N	sta C.C.,	orrell J.), Heimar	Karpen C
PRELIMINARY; (Tremblrel.	(TrEMBLrel. (TrEMBLrel.	nelanogast	detazoa; A dopterygo	, Drosophi 1227;	OM N.A.	96006; Pub	Celniker P.G., Sche	Lewis S.	Wortman	yle C., E	Agbayani Ragu A	Benos P.	Botchan M	Cawley S	., Delcher Joup L.E.,	Evangeli	Song F., G	Harvey I	Kalush F., Kodira C
T 1 5 09V3W5 001-MAY-2000	01-UCT-2002 (17E 01-MAR-2003 (TrE CG14052 protein.	CG14052. Drosophila melanogaster (Fruit fly).	ukaryota; l eoptera; E	Ephydroidea; Drosophilidae; Drosophila. NCBI TaxID=7227;	SEQUENCE FROM N.A.	MEDLINE=20196006; PubMed=10731132;	dams M.D., manatides)	eorge R.A.	utton G.G.	an K.H., De	bril J.F.,	Beeson K.Y.,	orkova D., urtis K.C.	herry J.M.	e Pablos B odson K)	Durbin K.J.	Ä	Harris N.L.	
RESULT Q9V3W5 ID Q AC Q DT 0		S S S	8 8 E	0 X X			RA RA RA		RA 60		RA PA		RA RA B	_	5 5 5 5 5 5		. •	RA H H	•

Hypothetical protein

```
ö
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Allerica B., McIntced M.P., McPherson D., Liu X., Mattels B., McIntced M.P., McPherson D., RA McIntolov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., RA Palazzolo M.P., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun H., Spier B. Spradling A.C., Stapleton M., Strong R., Sun E., RA Suirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J., Anilams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao, M., Walsenbach J., Sheng X.H., Zhong F.N., Zhan M., Zhang G., Zhao Q., Zheng L., RA Gibbs R.A., Myers B.W., Rubin G.M., Venter J.C.; Shu X., Smith H.O., Schence 287:2185-2195(2000).
                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.

Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,

Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,

Brans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,

Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,

Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,

Carlson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,

A postalez S., Frise E., Galle R.A., Hostin D., Howland T.J.,

Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,

I begwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,

Andintosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,

A poleb J., Paragas V., Park S., Patel S., Pfeiffer B.,

Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,

Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,

Milliams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;

Sequencing of Drosophila melanogaster genome.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROW N.A. Matthews B.B., Bayraktaroglu L., Campbell K., Histas S., Crosby M.A., Matthews B.B., Prochnik S.E., Smith C.D., Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D., Tupy J.L., Bergman C., Berman B., Carlson J.W., Cehriker S.E., Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N., Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E., Santh E., Shu S., Smutniak F., Whitfield E., Sabburner M., Gelbart W.M., Rubin G.M., Mangall C.J., Lewis S.E., Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 186;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              71.3%; Score 97; . DB 5; Length 186
68.4%; Pred. No. 1.8e-05;
ive 1; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
Submit AE003423; AAF45754.2; -.
FlyBase; FBGN0029606; CG14052.
SEQUENCE 186 AA; 21848 MW; 9BF94586AD4EFAD9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 3384 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   118 KHDHHDHHDHHDHHEH 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 КНКНИКНИКНИКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (TrEMBLrel. 23, (TrEMBLrel. 23, (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Adams M.D., Celniker
Submitted (MAR-2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q81L42;
01-MAR-2003 (
01-MAR-2003 (
01-MAR-2003 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q8IL42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q81L42
ID Q8
AC Q8
DT 01
DT 01
                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
```

```
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA MEDLINE-20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Hoskins R.A., Galle R.F.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazelj R.G., Champe M., Pfeiffer B.D.,
RA Brandon R.C., Raxer E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Ballew R.M., Basu A., Baxendale J., Bayrakteroglu L., Beagley E.M.,
RA Ballew R.M., Basu A., Bouck J., Bayrakteroglu L., Basalley E.M.,
RA Berman B.P., Bhandari D., Bolshakov S.,
RA Burtis K.C., Busam D.A., Buller H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
Charty J.M., Evangeliste C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Durbin K.J., Evangeliste C.C., Ferraz C., Ferriera S., Fleischmann M.,
ROGOR F., Gong F., Gorrell J.H., Gu Z., Glabat W.M., Glasser K.,
RA Harris N.L., Harvey D., Heimann T.J., Wei M.-H., Ibegwam C.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
                                                                                                                                                                                                                                                                                             Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W., Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K., Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S., Chan M., Sutherford K., Salzberg S.L., Craig A., Kyes S., Pertea M., Allen J., Selajom J., Suh B., Peterson J., Angluoli S., Pertea M., Allen J., Selajom J., Haft D., Mather M.W., Vaidya A.B., Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A., Worfer J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W., Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W., Fraser C.M., Barrell B., Linner M. M., Mandall C., Davis R.W.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CG4668 protein (LD30602p).
CG31738 OR CG4668 OR CG13261.
Drosophila melanogaeter (Fruit fly).
Bukaryota, Metazoa, Arthroda; Hexapoda; Insecta; Pterygota;
Bupotera; Endopterrygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
                          PF14_0407.
Plasmodium falciparum (isolate 3D7).
Eukaryota, Alveolata, Apicomplexa, Haemosporida, Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 97; DB 5; Length 3384;
Pred. No. 0.00025;
1; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Genome sequence of the human malaria parasite Plasmodium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL, AE014822; AAN37020.1; -.
Hypothetical protein.
SEQUENCE 3384 AA; 404576 MW; 340A187BCAEBAGFD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01.MAY-2000 (TrEMBLrel. 13, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 1413 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1; Mismatches
                                                                                                                                                                                                                                                                                  MEDLINE=22255705; PubMed=12368864;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      647 ЕНИННИННИННИН 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 кнининининининин 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    71.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 63.2
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nature 419:498-511(2002)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                             NCBI_TaxID=36329;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  falciparum.
                                                                                                                                                                                                                                                       STRAIN=3D7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9VJJ8
Q9VJJ8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  877760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOO OOS SOO OO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

```
;
0
Lasko P., Lei Y., Leviteky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Amerwlov G., Milshina N.V., Muzhy C., Moris J., Moshrefi A.,
Merwlov G., Milshina N.V., Nurphy L., Muzhy D.M., Nelson D.L.,
Nolson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Resse M.G.,
Reinert K., Remington K., Sampson M., Skupski M.P., Smith T.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Shue B.C., Stapleton M., Strong R., Smith T.,
Shue B.C., Wassarman D.A., Weinstock G.M., Weissenbach J.,
Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou S., Zhao Q., Zhao G., Zhao X., And Sheng X.H., Rong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers B.W., Rubin G.M., Venter J.C.;
"The genome sequence of Drosophila melanogaster.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Word, Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Miklos G.L.G.
                                                                                                                                                                                                                                                              STRAIN-Berkeley;
Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
Stapleton M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CG31738 protein.
CG31738 OR CG4668 OR CG13261.
Drosophila melanogaeter (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Boyaryota; Brachycera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 1413;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       69.1%; Score 94; DB 5; Length 141
63.2%; Pred. No. 0.00027;
ive 2; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                            Submitted (Apr-2002) to the EMBL/GenBank/DDBJ databases.
EMBL; AE003652; AAF53547.2; -.
EMBL; AX094813; AAM11166.1; -.
Flybase, FBGN0051738; CS31738.
Interpro; IPR003962; FNIII subd.
Interpro; IPR003961; FN III.
Interpro; IPR001865; Ribosomal_S2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1413 AA; 153288 MW; 93390C986EB40B03 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09VJJ9;
0-MX-2000 (TrEMBLrel. 13, Created)
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 1700 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROSITE; PS00962; RIBOSOMAL S2 1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN=Berkeley;
MEDLINE=20196006; PubMed=10731132;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1321 QHQHHMHSHHMHAHPH 1339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 КНКНИКНИКНИКНИКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRINTS; PR00014; FNTYPEIII.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 63.2
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF00041; fn3; 9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SM00060; FN3; 9
                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q9VJJ9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 4
       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
Ballow R. W. Badavan R. W. Badavan, B. Badadera P. Badadera P. B. Ballow R. M. Badavan R. Badavan R
```

.. 0

RESULT 5

S 셤

SERVE

```
Arabidopsis thaliana (Mouse-ear cress). Bubryophyta; Tracheophyta; Sukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN=cv. Columbia;

STRAIN=cv. Columbia;

Lin X., Kaal S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,

Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,

Bowman C.L., White O., Nierman W.C., Fraser C.M.;

Bowman C.L., White O., Nierman W.C., Fraser C.M.;

Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.

-I. SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

EMBL, AC011708; AAF19568 1;

-InterPro; IPRO0191, Znf_ring.

Pfam; PF00097; zf-C3H44; 1.

SWART; SM00184; RING; 1.
Force A., Amores A., Postlethwait J.;

"Hox cluster organization in the jawless vertebrate, Petromyzon marinus, and the evolution of the vertebrate Hox clusters.";

Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.

EMBL; AF414665; AAL61641.1; -.

REMBL; AF414665; AAL61641.1; -.

RICEPTO, IPRO01356; Homeobox.

Prom, PRO0046; homeobox.

REMBL; PRO0044; HOMEOBOX.

REMBL; PRO0010; Homeobox.

REMBL; PRO0010; Homeobox.

REMBL; PRO0010; Homeobox.

REMBL; PRO0011; HOMEOBOX.

REMBL; PRO0011; HOMEOBOX.

REMBL; PRO011; HOMEOBOX.

REMBL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                66.9%; Score 91; DB 10; Length 684; 66.7%; Pred. No. 0.00033; ive 0; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    67.6%; Score 92; DB 13; Length 404; 63.2%; Pred. No. 0.00015; arive 1; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 946203A42A7399B1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Putative RING zinc finger protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q8T029 PRELIMINARY; PRT; 782 AA.
Q8T029; 010703-2002 (TrEMBLrel. 21, Created)
01-UTM-2002 (TrEMBLrel. 21, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     782 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               684 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROSITE; PS50089; ZF RING 2; 1.
Metal-binding; Zinc; Zinc-finger.
SEQUENCE 684 AA; 76659 MW; 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 520 нинининининин 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2 НКННКНИКНИКНК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 нкинкинкинки 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity 66...
Local 2; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity 63.2
nes 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9SG87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 8
Q8T029
ID Q8T0
AC Q8T0
DT 01-J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9SG87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN=ATCC 15692 / PAO1;
STRAIN=ATCC 15692 / PAO1;
MEDLINE=20437337; PubMed=10984043;
MEDLINE=20437337; PubMed=10984043;
MEDLINE=20437337; PubMed=10984043;
Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P., Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., Galzer S.N., Folger K.R., Kas A., Larbig K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
"Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen.";
Nature 406.959-964(2000).
                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Petromyzon marinus (Sea lamprey).

Bukaryota, Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;

Petromyzontiformes; Petromyzontidae; Petromyzon.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
NCBI_TaxID=287;
                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       67.6%; Score 92; DB 16; Length 326; 63.2%; Pred. No. 0.00013; Live 1; Mismatches 6; Indels
                                                                                                                                                                                                                            Query Match
Best Local Similarity 63.2%; Pred. No. 0.00031;
Matches 12; Conservative 2; Mismatches 5; Indels
                                                                                                                                                                 Repeat.
SEQUENCE 1700 AA; 183716 MW; 4B82318D37AAD052 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        326 AA; 35772 MW; F948E09ACC2DB92A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q91447;
01-WAR-2001 (TrEMBLrel. 16, Created)
01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
01-0CT-2002 (TrEMBLrel. 22, Last annotation update)
probable metal transporter.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                326 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q8UVJ4;
01-WAR-2002 (TrEMBLrel. 20, Created)
01-WAR-2002 (TrEMBLrel. 20, Last seq.
01-WAR-2003 (TrEMBLrel. 23, Last ann
Hoxlw (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; AE004559; AAG04686.1; -.
INICEPTO; INFO02254; Catton_efflux.
Pfam; PF01545; Catton_efflux; 1.
PIGRFAMS; TIGR01297; CDF; 1.
                                         Pfam; PF00041; fn3; 9.
PRINTS; PR00014; FNYYBEIII.
SMART; SM00060; FN3; 9.
PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
                                                                                                                                                                                                                                                                                                                                                                                  1608 QHQHHWHSHHWHAHPH 1626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 163 нининининининини 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 НКННКННКННКНКНК 20
                                                                                                                                                                                                                                                                                                                                                 1 КНКНИКНИКНИКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity 63.2 tes 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Complete proteome.
SEQUENCE 326 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q8UVJ4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      091447
```

ö

RESULT 6

원 ò

080VJ4 11D 08 AC 08 AC 08 DJT 011 DJT 011 DJT 011 DJC DC DC OC DC DC OX NC

```
PRT; 1291 AA
                                        Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=20196006; PubMed=10731132;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   539 НОННЕННИННИННЫ 556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2 нкникникникн 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   66.9%;
                                      01-NOV-1998 (TrEMBLrel. 08, 01-NOV-1999 (TrEMBLrel. 12, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity 66.7
    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                    01-NOV-1998 (TrEMBLrel. 01-NOV-1999 (TrEMBLrel.
                                                                                               EG:115C2.10 protein.
EG:115C2.10 OR CG13363.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMART; SM00317; SET; 1
PROSITE; PS50280; SET;
SEQUENCE 1291 AA; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE PROM N.A.
                                                                                                                                                                                                                      NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN=BERKELEY;
                                                                                                                                                                                                                                                                                               Glover D.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9W5E0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q9W5E0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
    ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. Stableton N.A. Stableton M., Broketein P., Hong L., Agbayani A., Carlson J., Clampe M., Broketein P., Hong L., Agbayani A., Carlson J., Frise E., George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G., Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S., Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.B., Rubin G.M.,
                                                                                                                                                                                         STRAIN=Berkeley;
Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
Champe M., Chavez C., Dorsett V., Farfan D., Frise B., George R.,
Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
Yu C., Lewis S.B., Rubin G.M., Celniker S.,
Submitted (DEC-201) to the EMBL/GenBank/DDBJ databases.
-1- SIMILARITY: CONTAINS 1 SET DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                   Eukaryota, Metazoa, Arthropoda, Hexapoda, Insecta, Pterygota, Neoptera, Endopterygota, Diptera, Brachycera, Muscomorpha, Ephydroidea, Drosophilidae, Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Drosophila melanogaster (Fruit fly).
Eukaryota, Metazoa, Arthropoda, Hexapoda, Insecta, Pterygota,
Neoptera, Endopterygota, Diptera, Brachycera, Muscomorpha,
Ephydroidea, Drosophilidae, Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               66.9%; Score 91; DB 5; Length 782;
66.7%; Pred. No. 0.00037;
tive 1; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                66.9%; Score 91; DB 5; Length 825; 66.7%; Pred. No. 0.00039; tive 1; Mismatches 5; Indela
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C32D65359885B969 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        782 AA; 83863 MW; A7BDAD5DD44DE490 CRC64;
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     825 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                                                                                                                                                        FlyBase, FBgn0047197; BcDNA:LD36415.
InterPro; IPR001214; SET.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                           Drosophila melanogaster (Fruit fly)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL, AY089517; AAL90255.1; -
FlyBase; FBgn0025639; EG:115C2.10.
SEQUENCE 825 AA; 86859 MW; C32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           539 HQHHPHHHHHHHHHH 556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 НИНИНИНИНИНИНИ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 НКНИКНИКНИКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JUN-2002 (TrEMBLrel. 21, 01-JUN-2002 (TrEMBLrel. 21, 01-OCT-2002 (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     нони внининининини
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 66.9
Best Local Similarity 66.7
Matches 12, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 66.7
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART, SM00317, SET, 1.
PROSITE, PS50280, SET, 1
SEQUENCE 782 AA, 8386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EG:115C2.10 OR CG13363.
                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00856; SET; 1.
                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=7227;
                                        BCDNA: LD36415
```

QBT3U6

RESULT 9

ઠે 셤 Query Match

ઠે

RESULT 10

```
Adams M.D., Celniker S.E., Holf R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.P., Goorge R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandaell M.D., Zhang O., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D., Man K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Mikklos G.L.G., Abril J.F., Agbayani A., An H.-J., Bayraktaroglu L., Beasley E.M., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P., Abrits K.C., Busam D.A., Butle K.C., Busam D.A., Butle K.C., Davenport L.B., Davies P., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., Ab Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Pleischmann W., Ab Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                "Sequencing the distal X chromosome of Drosophila melanogaster."; Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
Salles C., Valenti P., Darlamitsou A., Henderson N., Campbell L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MX-2000 (TrEMBLrel. 13, Created)
01-MX-2000 (TrEMBLrel. 13, Last sequence update)
01-MX-2000 (TrEMBLrel. 23, Last annotation update)
01-MX-2003 (TrEMBLrel. 23, Last annotation update)
01-MX-2000 (TrEMBLrel. 23, Last annotation update)
EG:115C2.10 protein.
EG:115C2.10 OR CG13363.
Drosophila melanogaster (Fruit fly).
Brosophila melanogaster (Fruit fly).
Eghydroidea, Edopterygota; Diptera; Brachycera; Muscomorpha;
NCBI_TaxID=7227;
                                                                                                    Drosophila melanogaster (Fruit fly).
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 1291;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1291 AA; 137366 MW; 021DD566196196BA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
-!- SIMILARITY: CONTAINS 1 SET DOWAIN.
EMBL, AL031581; CAA20894.2; -..
FlyBase; FBGN0025639; EG:115C2.10.
InterPro; IRR001214; SET.
Pfam; PF00856; SET; 1.
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 91; DB 5; 1
Pred. No. 0.00058;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 1300 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1, Mismatches
```

407 AA.

PRELIMINARY;

```
SEQUENCE FROM N.A.

STRAIN=CV. Columbia,

MEDLINE=20083487; PubMed=10617197;

Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,

Lin X., Kaul S., Rounsley S.D., Shea T.P., Bernstead M.E., Feldblyum T.V.,

Fujii C.Y., Mason T.M., Dowman C.L., Barnstead M.E., Feldblyum T.V.,

Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,

Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,

Adams M.D., Carrera A.J., Crassy T.H., Goodman H.M., Somerville C.R.,

Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,

Salzberg S.L., Fraser C.M., Venter J.C.;

Sequence and analysis of chromosome 2 of the plant Arabidopsis
                                                                                                                                                                                                              Arabidopsis thaliana (Mouse-ear cress).

Bukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                           01-MAY-1999 (TrEMBLrel. 10, Created)
01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
                                                                                                                                                                                                        Putative pectinesterase
                                                                                      0920A3;
01-MAY-1999
                                                             0920A3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9VKD8
                                 09 ZQA3
                                                                                         SO DRAW REPART OF THE PROPERTY OF THE PROPERTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
Posler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., and Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris M.L., Harrardez J.R., Houck J., Harris M.L., Harrardez J.R., Houck J., A Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., And Jalai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Rasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Andrei B.E., Kodira C.D., Kraft C., McLeod M.P., McPherson D., Andrei B., McIntosh T.C., McLeod M.P., McPherson D., Andrei B., McIntosh T.C., McLeod M.P., McPherson D., Andrei B., McIntosh T.C., McLeod M.P., McPherson D., Rak Merkulov G., Milbhina N.V., Mobarry C., Morris J., Noshrei A., Nakon K., Nakon K., Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reses M.G., Rainert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Rainert K., Takon E., Strong R., Siden-Kamos I., Simpson M., Skrong R., Sun E., Shen H., Ray Spier E., Spradling A.C., Stapleton M., Skrong R., Wang A.H., Wang X., Ray Shi R. F., Zaveri J. S., Zhan M., Wang S., Yao Q.A., Ray Shen S., Woodage T., Worley K.C., Wu D., Yang S., Yao Q., Zheng X. H., Zhong F.N., Zhong W., Zhong K., Zhao Q., Zhao Q., Zheng X., Myes B.W., Rubin G.M., Venter J.C.; Zhu X., Smith H.O., R. Arney S., Wang S., Xao Q. M., Williams S.M., Myers B.W., Rubin G.M., Venter J.C.; Sho X., Smith H.O., R. Science 287:2185-2195(2000).

R. Science 287:2185-2195(2000).

R. Schence Bedyndes F.G.,1157-710.

R. Berly, REDOMAINS 1 SET DOWAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Glockner G., Eichinger L., Szafranski K., Pachebat J., Dear P., Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K., Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.; Sequence and Analysis of Chromosome 2 of Dictyostellum."; submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases. EMBL, AAM44363.1; -. Hypothetical protein. 233 AA, 20937 MW, AD4C7A2F86BB14B9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 1300;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        66.2%; Score 90; DB 5; Length 233; 66.7%; Pred. No. 0.00017;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        66.9%; Score 91; DB 5; Length 130
66.7%; Pred. No. 0.00059;
ive 1; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1300 AA; 137524 MW; 7186D0198CC16926 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Dictyostellum discoideum (Slime mold).
Eukaryota, Mycetozoa, Dictyostellida, Dictyostellum.
NCBI_TaxID=44689;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-0CT-2002 (TrEMBLrel. 22, Created)
01-0CT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Hypothetical protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     233 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FlyBase; FBGT0025639; EG:115C2.10.
Interpro; IPR001214; SET.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     539 НОННЕННИНННИН 556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 НХНИКНИКНИКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 НКНИКНИКНИКН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity 66.7
12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               12, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART; SM00317; SET; 1.
PROSITE; PS50280; SET;
SEQUENCE 1300 AA; 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF00856; SET; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=AX4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q8MP30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q8MP30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q8MP30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AC
DT
DT
DT
DT
DT
DT
RN
RN
RR
RA
RR
RR
SQ
SQ
SQ
                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ઠે
```

```
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota, Metazoa, Arthropoda, Hexapoda, Insecta, Pterygota, Bukoptera, Endoptera, Diptera, Brachycera, Muscomorpha, Ephydroidea, Drosophilidae, Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 66.2%; Score 90; DB 10; Length 407; 66.7%; Pred. No. 0.00027; ive 1; Mismatches 5; Indels
                                                                                                                                                                                                           Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
EMBL, AC06282; AAD20147.1; --
InterPro; IPR00070; Pectinesterase.
Pfam; PP01095; Pectinesterase; 1.
PROSITE; PS00503; PECTINESTERASE.2; 1.
SEQUENCE 407 AA; 44973 MW; DBAB7B8E166D0B42 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Drosophila melanogaster (Fruit fly)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN=Berkeley;
MEDLINE=20196006; PubMed=10731132;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       нинининининининин 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAY-2000 (TrEMBLrel. 13,
01-MAY-2000 (TrEMBLrel. 13,
01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 66.7
nes 12, Conservative
Nature 402:761-768(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                       SEQUENCE FROM N.A.
STRAIN=CV. Columbia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CG6686 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09VKD8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CG6686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ID DIT ACCOUNTS ACCOU
```

71 НИННИННИННИННН

ઠે ద ;

```
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beaaley E.M., Respon K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Rackotan M.R., Bouck J., Broketain P., Bolshakov S., Burtis R.C., Busam D.A., Butler H., Gadieu B., Center A., Chandra I., RA Burtis K.C., Busam D.A., Butler H., Gadieu B., Center A., Chandra I., RA Dedson K., Doug D. B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., RA Poels C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K., RA Foels C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K., RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., RA Harris N.L., Harvey D., Heiman T.J., Wernandez J.R., Houck J., RA Harris N.L., Harvey D., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Jalai M., Kalush F., Karpen G.H., Ke Z., Kennison J., Lai Z., Lin X., Martei B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lin X., Martei B. McIntosh T.C., McLeod M.P., McPherson D., Lai Z., Linag Y., Lin X., Markei B. McIntosh T.C., McLeod M.P., Nebreson D., McBherson D. M., Nelson K.A., Nixon N., Nausny D.M., Nelson D.L., Ralazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Rahner K., Remington K., Saunders R.D.C., Scheels F., Shen H., Rayleskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang S., Turner R., Venter E., Wang S., Yao Q.A., Rytskas R., Tector C., Turner R., Venter E., Wang S., Yao Q.A., Rytskas R., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Rytskas R., Myers B.W., Rubin G.M., Venter J.C.; Rhy Gibbs R.A., Myers B.W., Rubin G.W., Venter J.C.; Rhy Gibbs R.A., Myers B.W., Rubin G.W., Venter J.C.; Rhy Gibbs R.A., Myers B.W., Rubin G.W., Venter J.C.; Rhy Gibbs R.A., Myers B.W., Rubin G.W., Venter J.C.; Rhy Gibbs R.A., Myers B.W., Rubin G.W., Venter J.C.; Rhy Gibbs R.A., Myers B.W., Rubin G.W., Venter J.C.; Rhy Gibbs R.A., Myers B.W., Rubin G.W., Venter J.C.; Rhy Gibbs R.A., Myers B.W., Rubin G.W., Zhong Y. Shith H. P., Schong Y. W. Saith H. P., S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Celinier S. E., Adams M. D., Kronmiller B., Wan K.H., Holt R.A.,
A Bvans C.A., Gocayne J.D., Amanatides P.G., Breandon R.C., Rogers Y.,
Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
A Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
A McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,
A McIntosh T.C., Moy M., Park S., Patel S., Pfeiffer B., Scheeler F.,
A Pacleb J., Paragas V., Park S., Patel S., Fleiffer B., Scheeler F.,
Williams S.M., Savirskas R., Tector C., Tyler D.,
Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
Sequencing of Drosophila melanogaster genome.";
Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mista S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K., Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D., Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E., Kronmiller B., Marshall B., Millburn G., Richter J., Russo S., Searle S.M.J., Smith B., Millburn G., Richter J., Russo S., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.; Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.; Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PlyBase;
Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
EMBL; AE003634; AAN10790.1; -.
PlyBase; FBGN0032389; CG6686.
InterPro; IPR005011; SART_1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0F379483CEFD5D85 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF03343; SART-1; 1. SEQUENCE 964 AA; 111902 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Science 287:2185-2195(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
```

```
RX PEQUENCE FROM N.A.

RX MEDLINE-2019-6006; PubMed=10731132;

RADIANE-2019-6006; PubMed=10731132;

RADIANE-2019-6006; PubMed=10731132;

RADIANE-2019-6006; PubMed=10731132;

RADIANE-2019-6006; PubMed=10731132;

RADIANE-2019-6006; Mortman J. R., Yandell M.D., Zhang Q., Chen L.X., Button G.G., Mortman J. R., Yandell M.D., Change M., Pefelfer B.D., Man K.H., Doyle C., Badrer E.G., Helt G., Nelson C.R., Gabor G.L., Man K.H., Doyle C., Badrer E.G., Helt G., Nelson C.R., Gabor G.L., Man K.H., Doyle C., Barden B.D., Barden B.D., Barden M., Basul M., Basul M., Basul M., Basul M., Basul M., Bayaraktaroglu L., Beasley E.M., Ballew R.M., Gabor B.D., Buttis R., Bouck J., Brokketen P., Brottler P., Wang S., Poller P., Brottler P., Wang S., Poller P., Brottler P.
                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J., Champe M., Chavez C., Dorsett V., Farfan D., Frise B., George R., Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K., Yu C., Lewis S.B., Rubin G.M., Celniker S.; Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
                                            ï
                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Metazoa; Arthropoda; Haxapoda; Insecta; Pterygota; Neoptera; Endoptera; Endoptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
  Length 964;
                                            Indels
                                                                                                                                                                                                                                                                          01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
65.8%; Score 89.5; DB 5; 66.7%; Pred. No. 0.00069;
                                                                                                                                                                                                                                      970 AA.
                                          3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                  Drosophila melanogaster (Fruit fly).
                                                                                                                                                                                                                                      PRT;
                                                                                  1 КНКНИКНИКНИКН-НКНКНК 20
                                                                                                            66.78;
                                          14; Conservative
                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                            LD23187p (CG6686-PA).
                   Best Local Similarity
Matches 14; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-Berkeley;
  Query Match
                                                                                                                                                                                                                                    Q95TU2
Q95TU2;
                                                                                                                                                                                            RESULT 15
Q95TU2
                                                                                                                            셤
                                                                                                                                                                                                                                                          ð
```

```
Celliker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,

Geniker S.E., Adams M.D., Amanatides P.G., Brandon R.C., Rogers Y.,

By Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,

Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busman D.A.,

Banzon J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,

Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,

Rerriera S., Frise E., Galle R.F., Gargy N.S., George R.A.,

Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,

Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,

Rolltosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,

McIntosh T.C., Moy M., Wurphy S., Petel S., Pétiffer B.,

Pacleb J., Paragas V., Park S., Patel S., Pétiffer B.,

Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,

Rippleton M., Strong R., Svirskes R., Tector C., Tyler D.,

Rilliams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;

"Sequencing of Drosophila melanogaster genome.",

"Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
Maitze S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
Maitze S., Crosby M.A., Matthews J.S., Prochnik S.E., Smith C.D.,
Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
Submitted (MAR-2000) to the EMBL/Genbank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A. Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.; Adams M.D., Celniker S.E., Gibbs R.A., Rubin databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 970;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
EMBL; AYOS525; AAL13754.1; -.
EMBL; AE003634; AAF53138.2; -.
FlyBases; FBG10032388; CG6686.
InterPro; IFROGO011; SART_1.
Fign, PF03343; SART-1; 1.
SEQUENCE 970 AA; 112575 WW; A8C93B37F9B89228 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 65.8%; Score 89.5; DB 5; Best Local Similarity 66.7%; Pred. No. 0.00069; Matches 14; Conservative 3; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
[3]
SEQUENCE FROM N.A.
```

Search completed: January 20, 2004, 18:26:24 Job time : 27.6377 secs 셤

1 КНКНИКНИКН-НКНИКК 20

ઠે

Gaps

ä

3; Indels

```
January 20, 2004, 18:05:34; Search time 31.0145 Seconds (without alignments) 102.356 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 the number of results predicted by chance to have a r than or equal to the score of the result being printed, of by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A Geneseq 197un03:*

| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AAA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AAA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1107863
                           GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1107863 segs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                         OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                              US-10-018-103A-6
136
1 KKHHHKHHHKKHHKK 20
                                                                                                                                                                                                                                                                                                                                                                                                                                     BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. is the n
score greater than
and is derived by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Post-processing:
                                                                                                                                                                                                                                                                                                              Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                        Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Database
                                                                                                                                                                                                Run on:
```

Description	Histidine copolyme	Human liver peptid	Peptide #4697 enco	Peptide #4801 enco	Protein #4583 enco	Human brain expres	Human bone marrow	Peptide #4697 enco	Peptide #4792 enco
ID		ABG52113	ABB32046	ABB37295	ABB22584	AAM57997	AAM70433	AAM18263	AAM30755
80	22	22	22	22	22	22	22	22	22
% Query re Match Length DB I	20	49	49	4	49	49	4	49	49
% Query Match	100.0	71.3	71.3	71.3	71.3	71.3	71.3	71.3	71.3
Score	136	97	97	97	97	97	97	97	97
Result No.	-	7	٣	4	ស	9	7	ω	σ

Peptide #4559 enco Novel human diagno Human liver peptid Peptide #4866 enco Peptide #4866 enco Peptide #4866 enco Peptide #4815 enco Human bone marrow Peptide #4815 enco Novel human diagno HIV A59P protein s HIV A59P protein s HIV A59P protein s HIV A59P protein s HIV G75A protein s	AAE06231 standard; peptide; 20 AA. AAE06231; 25-SEP-2001 (first entry) Histidine copolymer; enzyme deficiency; phenylalanine hydroxylase; ornithine transcarbamylase; adenosine deaminase; tyrosine hydroxylase; ornithine transcarbamylase; alphal-antitrypsin; apolipoprotein E; chronic granulomatous disease; alphal-antitrypsin; apolipoprotein E; chronic granulomatous disease; sichle cell; beta-thalassasemia; anaemia; chronic granulomatous disease; sichle cell; beta-thalassasemia; anaemia; phenylketconuria; Paconi's anaemia; heemophilia; muscular dystrophy; cystic fibrosis; Parkinson's disease; retninits pigmentosa; cytostatic; lysosomal storage disease; mycopolysaccharide type I disease; cardiant; diabetic retinopathy; human immunodeficiency virus disease; cardiant; diabetic retinopathy; human immunodeficiency virus disease; infection; cardiac disease; peripheral vascular disease; antisickling; arthritis; unidentified. WO200147496-A1. 29-DEC-1999; 99US-0173576.
AAMOS 877 ABG47005 ABG47005 ABG471093 ABG52171 ABB37360 AAB37360 AAB37360 AAM70490 AAM70490 AAM70490 AAM70490 AAM70490 AAM65933 AAG64772 AAG64772 AAG99330 AAG99330 AAG99330 AAG99330 AAG72087 AAG99330 AAG72087 AAG99330 AAG993300 AAG993300 AAG993300 AAG993300 AAG993300 AAG99300	peptide. deficiency; deficien
222222222222222222222222222222222222222	le; 2 (K per an
22222222222222222222222222222222222222	standard; peptide; (i) (first entry) copolymer; HHH-K copolymer; enzyme transcarbamylase; transcarbamylase; transcarbamylase; transcarbamylase; transcarbamylase; copolymer; enzyme transcarbamylase; transcarbamylase; transcarbamylase; transcarbamylase; transcarbamylase; transcarbamylase; transcarbamylase; retinopathy; human lisease; peripheral i, haemostatic; virr fied. 196-Al. 100; 2000WO-US34603
	tandard; p (first copolymer; transcarba leoside ph anulomatou yapulomatou yapulomatou yapulomatou yapulomatou yapulomatou yapulomatou yapulomatou yapulomatou sease; per storage di etinopathy sease; per haemostat ed. (6-Al.
	stand i 001 (e copo e tran ucleos granul ibrosi ibrosi ibrosi retin diseas c; hae fied. 496-A1
00000000000000000000000000000000000000	AAE06231 sta AAE06231; 25-SEP-2001 Histidine co conithine tr purine mucle chronic gran familial hyp phenylketonu cystic fibro cystic cystic fibro cystic cysti
011111111111111111111111111111111111111	AABSOULT AAB

```
ABB32046;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                  Penn SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ABB3204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                    The invention relates to a pharmaceutical agent delivery composition comprising a transport polymer comprising a linear or branched peptide comprising a transport polymer comprising a linear or branched peptide having about 10 to 300 amino acid residues, having about 5 to 100% instidine residues, and optionally having 1 to 95% non-histidine residues, at least one pharmaceutical agent and optionally one or more residues, at least one pharmaceutical agent delivery composition is useful for polymer. The pharmaceutical agent delivery composition is useful for delivering a pharmaceutical agent delivery composition is useful for closed in gene therapy for treating adenosine deaminase deficiency, purine used in gene therapy for treating adenosine deaminase deficiency, nucleoside phosphorylase deficiency, chronic granulomatous disease with defective p47phox, sickle cell with HBS, beta-thalasseamla due to defective p47phox, sickle cell with HBS, beta-thalasseamla due to defective p47phox, sickle cell with HBS, beta-thalasseamla due to defective p47phox, sickle cell with HBS, beta-thalasseamla due to defective p47phox, sickle celliney, laphal-antitrypsin deficiency, to phenylalanine bydroxylase deficiency, Faconi's anaemia, apolipoprotein contituine transcarbamylase deficiency, retnitis pigmentosa, lysosomal storage tyrosine hydroxylase deficiency, retnitis pigmentosa, lysosomal storage tyrosine hydroxylase deficiency, retnitis pigmentosa, lysosomal storage clisease (i.e., mycopolysacenite type 1, Hunter, Hurler, Hurler, disease (i.e., mycopolysacenite type 1, Hunter, Hurler, disease, virus indection, acquired anaemia, cardiac and peripheral vascular disease condiabetic retinopathy, human cardiac and peripheral vascular disease infection, acquired anaemia, cardiac and peripheral vascular disease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                       Pharmaceutical composition useful for delivering therapeutic agent for treating diseases, comprises a peptide characterized with specified number of amino acids and a specified percentage of histidine residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
hypercholesterolaemia; coronary heart disease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0%; Score 136; DB 22;
100.0%; Pred. No. 3.4e-11;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Human liver peptide, SEQ ID No 30761.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 ККННИКНИКИНИКК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ABG52113 standard; Peptide; 49
                                                                                                                                                                   Claim 3; Page 36; 64pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000US-0180312.
2000US-0207456.
2000US-0608408.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30-JAN-2001; 2001WO-US00664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 100.
Matches 20; Conservative
                                                                  WPI; 2001-425579/45
            (MIXS/) MIXSON A J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO200157273-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        04-FEB-2000; 26-MAY-2000; 30-JUN-2000; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-FEB-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    39-AUG-2001
                                          Mixson AJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 2
ABG52113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    쉱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
```

```
The invention relates to a single exon nucleic acid probe (SENP) (I) for measuring human gene expression in a sample derived from human adult liver, comprising one of 13109 defined nucleotide sequences given in the liver complements/ fragments). The probe hybridises at high specification (or complements/ fragments). The probe hybridises at high stringency to a nucleic acid molecule expressed in the human adult extringency to a nucleic acid molecule expressed in the human adult liver. (I) may be used for predicting, measuring and displaying gene expression in samples derived from human adult liver. The gene service expression in samples derived from human and hypercholesterolaemia which hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which seasociated with coronary heart disease. ABG57318 + ABG59310 represent in man liver single exon encoded peptides of the invention.

Note: The sequence information for this patent does not appear in the printed specification but was obtained in electronic format directly from WIPO at ftp. WiDo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Peptide #4697 encoded by breast cell single exon nucleic acid probe.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
                                                                                                                                                                                                                                                                                                                                           Human genome-derived single exon nucleic acid probes useful for analysing gene expression in human adult liver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human; microarray; single exon probe; gene expression; breast;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 97; DB 22; Length 49;
Pred. No. 6.7e-06;
2; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 27; SEQ ID No 30761; 658pp; English
                                                                                                                                                                                                                          Chen W, Rank DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ABB32046 standard; Peptide; 49 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (MOLE-) MOLECULAR DYNAMICS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |||:||| ||||:
13 нингинининик 29
                                                                                                                                                                (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 нинкиникиник 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27-SEP-2000; 2000US-0236359.
04-OCT-2000; 2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            71.3%;
Local Similarity 70.6%;
hes 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     30-JAN-2001; 2001WO-US00662.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2000US-0632366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2000US-0234687
03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0236359.
04-OCT-2000; 2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                 Hanzel DK,
                                                                                                                                                                                                                                                                                           WPI; 2001-488898/53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        49 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           disease; cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200157271-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      04-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26-MAY-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21-SEP-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-FEB-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       09-AUG-2001.
```

```
The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart (see ABA21535-ABA41305). The present sequence is a protein encoded by one such probe. The probes may be used for predicting, measuring and displaying
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Protein #4583 encoded by probe for measuring heart cell gene expression.
                                                                                                                                     measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human fetal liver. The present sequence is a peptide encoded by a single exon nucleic acid probe of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Single exon nucleic acid probes for analyzing gene expression in human hearts -
                               Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human fetal liver -
                                                                                                                        The invention relates to a single exon nucleic acid probe for
                                                                                     Claim 27; SEQ ID NO 29930; 639pp + sequence listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human, gene expression, heart, microarray, vascular system, cardiovascular disease, hypertension, cardiac arrhythmia,
                                                                                                                                                                                                                                                                                                                                    71.3%; Score 97; DB 22; Length 49; 70.6%; Pred. No. 6.7e-06; ive 2; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 15; SEQ ID No 24354; 530pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 щ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Rank
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ABB22584 standard; Protein; 49 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (MOLE-) MOLECULAR DYNAMICS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                           3 ННИКИННККИНИКНИК 19
                                                                                                                                                                                                                                                                                                                                                                                                                                 13 HHRHHHHHHHHR 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0236359.
04-OCT-2000; 2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26-MAY-2000; 2000US-0207456.
30-JUN-2000; 2000US-0608408.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 30-JAN-2001; 2001WO-US00666.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    04-FEB-2000; 2000US-0180312
                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 70.00
The 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      congenital heart disease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2001-488899/53.
WPI; 2001-483447/52.
                                                                                                                                                                                                                                                                                                       49 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200157274-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23-JAN-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09-AUG-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABB22584;
                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Penn SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ABB22584
 #X##X%XCCCCCCCXX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                       from human breast and BT 474 cells. The method involves contacting the probes with a collection of detectably labelled nucleic acids derived from human breast and BT 474 cells. The method involves contacting the probes with a collection of detectably labelled nucleic acids derived from mRNA of human breast, and then measuring the label.

Concode proteins of the microarray. The probes are useful for verifying the expression of regions of genomic DNA predicted to encode proteins. They are useful for gene discovery, and for ecode proteins. They are useful for gene discovery, and for expression analysis is useful for assessing breast disease. Gene expression analysis is useful for assessing the toxicity of chemical agents on cells. The microarray of this invention presents a far greater diversity of probes for measuring gene expression, with far-less bias chan expressed sequence tag microarrays. The method is suitable for rapid production of functional information from genomic sequence. The present sequence is a peptide encoded by a single exon nucleic acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human; foetal liver; gene expression; single exon nucleic acid probe.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                            of single exon
                                                                  New spatially-addressable set of single exon nucleic acid probes, useful for measuring gene expression in sample derived from human breast, comprises number of single exon nucleic acid probes -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Peptide #4801 encoded by human foetal liver single exon probe.
                                                                                                                                         Claim 27; SEQ ID NO 15014; 327pp + sequence listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           71.3%; Score 97; DB 22; Length 49; 70.6%; Pred. No. 6.7e-06;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3; Indels
                                                                                                                                                                            The invention relates to a spatially-addressable set
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Chen W, Rank DR;
Rank DR:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ABB37295 standard; Peptide; 49 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (MOLE-) MOLECULAR DYNAMICS INC.
Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 ННИМИНИМИНИМ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |||:||| ||| |||:
нинкининининик 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26-MAY-2000, 2000US-0207456.
30-UJN-2000, 2000US-060B40B.
03-MG-2000, 2000US-063366.
21-SEP-2000, 2000US-0234687.
27-SEP-2000, 2000US-0236359.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30-JAN-2001; 2001WO-US00669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     04-FEB-2000; 2000US-0180312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               04-OCT-2000; 2000GB-0024263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                04-FEB-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                  probe of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hanzel DK,
Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity
nes 12; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         49 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200157277-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Seguence
Penn SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Penn SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Loca
Matches
ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

Gaps

Gaps

.. 0

Indels

٠<u>.</u>

```
The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human bone marrow. They can be used to measure gene expression in bone marrow samples, which may enable the improved diagnosis and treatment of cancers such as lymphoma, leukaemia and myeloma. The present sequence is a protein encoded by one of the probes of the invention.
                                                                                                                                                                                                                                    gene expression analysis; probe;
                                                                                                                                                                                                         Human bone marrow expressed probe encoded protein SEQ ID NO: 30739.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Human genome-derived single exon nucleic acid probes useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Example 4; SEQ ID NO: 30739; 658pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          71.3%; Score 97; DB 22; Length 49; 70.6%; Pred: No. 6.7e-06; 1.ve 2; Mismatches 3; Indels
                                                                                                                                                                                                                                                       cancer; leukaemia; lymphoma; myeloma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        analyzing gene expression in human bone marrow
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Rank DR;
                                                                                                                                                                                                                                            expressed exon;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAM18263 standard; Protein; 49 AA
                                                                                                                            AAM70433 standard; Protein; 49 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (MOLE-) MOLECULAR DYNAMICS INC
 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 НИНКНИНККИННКНИНК 19
                                                                                                                                                                                                                                                                                                                                                                                                             26-MAY-2000; 2000US-0207456.
30-JUN-2000; 2000US-0608408.
03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0234687.
04-OCT-2000; 2000GB-0024263.
                          з нинкиниккиники 19
                                                                                                                                                                                                                                                                                                                                                                     30-JAN-2001; 2001WO-US00668
                                                                                                                                                                                                                                                                                                                                                                                                 2000US-0180312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |||:||| ||| ||||
HHHRHHHHHHHHHR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                            |||:||| ||| |||:
13 HHHRHHHHHHHHHR
                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Penn SG, Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2001-488900/53
                                                                                                                                                                                                                                             bone marrow
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       49 AA;
                                                                                                                                                                                                                                                                                                                 WO200157276-A2.
                                                                                                                                                                                                                                                                                       Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             12-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                    04-FEB-2000;
                                                                                                                                                                                     06-NOV-2001
                                                                                                                                                                                                                                                             microarray;
                                                                                                                                                                                                                                                                                                                                             09-AUG-2001
    12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAM18263;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13
                                                                                                                                                           AAM70433;
                                                                                                                                                                                                                                             Human;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 8
    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAM18263
                                                                                                 RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5 X X X E X
                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                ð
                                                         임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is a protein encoded by one of the probes of the invention.
                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                      Human brain expressed single exon probe encoded protein SEQ ID NO: 30102.
     gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, monitoring and prognosing diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Single exon nucleic acid probes for analyzing gene expression in human
                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                     Human; brain expressed exon; gene expression analysis; probe;
microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
epilepsy; cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Example 4; SEQ ID NO: 30102; 650pp + Sequence Listing; English.
                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 49;
                                                                                                                                                              71.3%; Score 97; DB 22; Length 49; 70.6%; Pred. No. 6.7e-06; 1.ve 2; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 97; DB 22;
Pred. No. 6.7e-06;
                                                                                                            at ftp.wipo.int/pub/published_pct_sequences
                                                                                                                                                                                            Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Rank DR;
                                                                                                                                                                                                                                                                                                                       AAMS7997 standard; Protein; 49 AA.
                                                                                                                                                                                            5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Chen W,
                                                                                                                                                                                                                                                 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      71.3%;
70.6%;
                                                                                                                                                                                                                        з нинхиникиник 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0236359.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26-MAY-2000; 2000US-0207456.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2000GB-0024263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30-JAN-2001; 2001WO-US00667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2000US-0180312
                                                                                                                                                                                                                                       |||:||| ||| |||
HHHRHHHHHHHHR
                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Penn SG, Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2001-483446/52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                 Query Match
Best Local Similarity
Matches 12; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             49 AA;
                                                                                                                                        49 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200157275-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           04-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             04-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                05-NOV-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                      AAM57997;
                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             brains
                                                                                                                                                                                                                                                                                               RESULT 6
                                                                                                                                                                                                                                                                                                             AAM57997
                                                                                                                                                                                                                                                                                                                                                        셤
             ઠે
```

ö

Gapa

S

```
30-JAN-2001; 2001WO-US00663
                                                                                                                                                                                                                                                                                                 49 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO200157270-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           04-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09-AUG-2001.
                                                                                                                                                                                                                                                                                                  Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAM05877;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Penn SG,
                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                         Matches
 ઠ
                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                      (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded by one such probe. The SENPs are derived from human HeLa cells. The SENPs can be used to produce a single exon microarray, which can be used for measuring human gene expression in a sample derived from human cervical epithelial cells. By measuring gene expression, the probes are therefore useful in grading and/or staging of diseases of the cervix, notably
                                                                                                                                                                                                                                                                                                                                                                                                            Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Peptide #4792 encoded by probe for measuring placental gene expression.
                                                                                                                                                                                                                                                                                                                              The present invention relates to human single exon nucleic acid probes
Peptide #4697 encoded by probe for measuring cervical gene expression.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                    human; microarray; gene expression; cervical epithelial cell;
                                                                                                                                                                                                                                                                           Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human cervical epithelial cells -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         71.3%; Score 97; DB 22; Length 49; 70.6%; Pred. No. 6.7e-06;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score y,;
Pred. No. 6.7e-06;
Pred. .....hes 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      microarray; human; placenta; antenatal diagnosis;
                                                                                                                                                                                                                                                                                                           Claim 27; SEQ ID No 23089; 487pp; English
                                                                                                                                                                                                                                    Chen W, Rank DR,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAM30755 standard; Protein; 49 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2;
                                                                                                                                                                                                                (MOLE-) MOLECULAR DYNAMICS INC.
                                                                                                                                          26-MAY-2000; 2000US-0207456.
30-JUN-2000; 2000US-0608408.
03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-023559.
04-OCT-2000; 2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 ННИКНИНКИНИКИНИК 19
                                                                                                             30-JAN-2001; 2001WO-US00670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HHREHHERHERERERE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                17-OCT-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                    Hanzel DK,
                                                                                                                                                                                                                                                       WPI; 2001-488901/53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
nes 12; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                       49 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genetic disorder
                                                                                                                                                                                                                                                                                                                                                                                                    cervical cancer.
                             cervical cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          40200157272-A2.
                                                                     WO200157278-A2
                                                  Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens
                                                                                                                                 04-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              09-AUG-2001
                                                                                         09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                    Penn SG,
                    Probe;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
```

```
Probe; human; breast disease; breast cancer; development disorder; inflammatory disease; proliferative breast disease; non-carcinoma tumour.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The present invention relates to single exon nucleic acid probes (SENP: see AA13115-AA157546). The present sequence is a peptide encoded by one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of human genetic disorders.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Peptide #4559 encoded by probe for measuring breast gene expression.
                                                                                                                                                                                                                                                                                                                                                                                                                                          Human genome-derived single exon nucleic acid probes useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          71.3%; Score 97; DB 22; Length 49; 70.6%; Pred. No. 6.7e-06; Live 2; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       analyzing gene expression in human placenta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 27; SEQ ID No 31024; 654pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Chen W, Rank DR;
                                                                                                                                                                                                                                                                                                                 Chen W, Rank DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAM05877 standard; Protein; 49 AA.
                                                                                                                                                                                                                                                   (MOLE-) MOLECULAR DYNAMICS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 ННИКНИНККИНИКНИК 19
                                                                                    03-AUG-2000; 2000US-0632366.
21-SED-2000; 2000US-0234687.
27-SED-2000; 2000US-0236359.
04-OCT-2000; 2000GB-0024263.
                             26-MAY-2000; 2000US-0207456
30-JUN-2000; 2000US-0608408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  04-FEB-2000; 2000US-0180312.
26-MAY-2000; 2000US-0207456.
30-JUN-2000; 2000US-0608408.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0236359.
2000US-0180312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        29-JAN-2001; 2001WO-US00661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000GB-0024263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13 HHRHEREHHHRHR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12; Conservative
                                                                                                                                                                                                                                                                                                                 Penn SG, Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                               WPI; 2001-488897/53.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
```

```
WPI; 2002-114183/15.
                                      ABG40067;
                      Sequence
                          Aatches
                                  RESULT 11
                                   ABG40067
                                      ò
                              셤
```

```
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABG17093;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
셤
                                                                                                                   The present invention relates to novel single exon nucleic acid probes (see AA100010-AA110067). The present sequence is a peptide encoded by one such probe. The probes are useful for measuring human gene expression in a human breast sample, where the probe hybridises at high stringency to a nucleic acid expressed in the human breast. The probes are useful for predicting, diagnosing, grading, stading, monitoring and prognosing diseases of the human breast, particularly those diseases with polygenic aetiology. The diseases include: breast cancer, disorders of development, inflammatory diseases of the breast. fibrocystic changes, proliferative
                                                                                                                                                                                                                                                                                           printed
                                                                                                                                                                                                                                                                         breast disease and non-carcinoma tumours.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human peptide encoded by genome-derived single exon probe SEQ ID 29732.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human; single exon probe; asthma; lung cancer; COPD; ILD; chronic obstructive pulmonary disease; interstitial lung disease; familial idiopathic pulmonary fibrosis; neurofibromatosis; tuberous sclerosis; Gaucher's falsoase; Niemann-Pick disease; Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis; pulmonary haemosiderosis; lymphangioleiomyomtosis; Karagener syndrome; pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia; primary ciliary dyskinesis; pulmonary hypertension; hyaline membrane disease.
                                   Novel single exon nucleic acid probe used to measuring gene expression
                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                           Score 97; DB 22; Length 49;
Pred. No. 6.7e-06;
2; Mismatches 3; Indels
                                                                                          claim 27; SEQ ID No 14617; 322pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rank DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABG40067 standard; Peptide; 49 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (MOLE-) MOLECULAR DYNAMICS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Chen W,
                                                                                                                                                                                                                                                                                                                                                                                                                                                              з нинкиниккиник 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |||:||| |||||1||
13 нияхнинининия 29
                                                                                                                                                                                                                                                                                                                                                                                                  71.3%;
70.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30-JAN-2001; 2001WO-US00665.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2000US-207456P.
2000US-0608408.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21-SEP-2000; 2000US-234687P.
27-SEP-2000; 2000US-236359P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              04-OCT-2000; 2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               03-AUG-2000; 2000US-0632366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                  12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Penn SG, Hanzel DK,
        WPI; 2001-476286/51
                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                  49 AA;
                                                            in a human breast
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200186003-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19-AUG-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                04-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26-MAY-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  30-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15-NOV-2001.
```

```
The invention relates to a spatially-addressable set of single exon mucleic acid probes for measuring gene expression in a sample derived mucleic acid probes for measuring gene expression in a sample derived from human lung comprising single exon mucleic acid probes having one of 12614 mucleic acid eaquences mentioned in the specification, or their complements or the 12397 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes. Also included are a microarray comprising the novel set of probes. Also included are a microarray comprising the array with a nucleic acid expressed in the human lung measuring gene expression in a maple derived from human lung comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung comprising (a) algorithmically predicting at least one exon from genomic sequences (c) the eukaryote; and (b) detecting specific hybridisation of detectably the babel denoted from entaryote lung man, to a single exon probe, laving a fragment identical to the predicted exon, the probe is included comprising (a) identifying exons from genomic sequence by the method comprising (a) identifying exons from genomic sequence by the method comprising (a) identifying exons from genomic sequence by the method comprising a probe with the exon, where a common pattern of microarrays having a probe with the exon, where a common pattern of microarrays having a probe with the exon, where a common pattern of microarrays having a probe with the exon, where a common pattern of microarrays having a probe with the exon, where a common pattern of c expression of the exons in the tissues and/or cell types using frames for the exons should be assigned to a single gene of 12011 sequences, mentioned in the probes are used for gene of 12011 sequences, mentioned in the probes are used for gene of c such as each man, lung cancer, chronic observative pulmonary disease. (COPD), interstitial lung disease (ILD), familiar seq
Spatially-addressable set of single exon nucleic acid probes, used to measure gene expression in human lung samples
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         71.3%; Score 97; DB 23; Length 49; 70.6%; Pred. No. 6.7e-06; ive 2; Mismatches 3; Indels
                                                                                                                                    Claim 27; SEQ ID No 29732; 634pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Novel human diagnostic protein #17084.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ABG17093 standard; Protein; 52 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 ННИХНИНКИНИКИ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
les 12; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             49 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18-FEB-2002
```

```
The invention relates to a single exon nucleic acid probe (SENP) (I) for measuring human gene expression in a sample derived from human adult liver, comprising one of 13109 defined nucleotide sequences given in the specification (or complements/ fragments). The probe hybridises at high stringency to a nucleic acid molecule expressed in the human adult liver stringency to a nucleic acid molecule expressed in the human adult capression in samples derived from human adult liver. The genes capression in samples derived from human adult liver. The genes in the propertipoproteinsemia, hyperlipidaemia and hypercholesterolaemia which is associated with coronary heart disease. ABG47348-ABG5930 represent human liver single exon encoded peptides of the invention.

The information for this patent does not appear in the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Peptide #4759 encoded by breast cell single exon nucleic acid probe.
                                                                                                                                                                                                                                                                                                                                                                                                                       Human genome-derived single exon nucleic acid probes useful for analysing gene expression in human adult liver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human; microarray; single exon probe; gene expression; breast;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 97; DB 22; Length 292;
Pred. No. 4e-05;
2; Mismatches 4; Indel8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 27; SEQ ID No 30819; 658pp; English
                                                                                                                                                                                                                                                                                                                            Chen W, Rank DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABB32108 standard; Peptide; 292 AA.
                                                                                                                                                                                                                                                                                 (MOLE-) MOLECULAR DYNAMICS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            нинкиннккиннкк 20
                                                                                          2000US-0180312.
2000US-0207456.
2000US-0608408.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 71.3%;
Best Local Similarity 66.7%;
Matches 12; Conservative
                                                                                                                                                       03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0236359.
                                                                                                                                                                                                                                     04-OCT-2000; 2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ||| ||| ||| ||| ||| 81 HHHPHHHHHPRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            04-FEB-2000; 2000US-0180312.
26-MAY-2000; 2000US-0207456.
30-JUN-2000; 2000US-0608408.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30-JAN-2001; 2001WO-US00662.
                                           30-JAN-2001; 2001WO-US00664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-FEB-2002 (first entry)
                                                                                                                                                                                                                                                                                                                            Penn SG, Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                           WPI; 2001-488898/53.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   292 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              disease; cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200157271-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens
                                                                                          04-FEB-2000;
26-MAY-2000;
                                                                                                                                        30-JUN-2000;
09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABB32108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABB32108
  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The invention relates to isolated polynuclectide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PGR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The cand gene mapping, and in recombinant production of (II). The polymuclectides are also used in diagnostics as expressed sequence tags correstor normal activity of (II) or to treat disease states involving to restore normal activity of (II) or to treat disease states involving correction or samplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating clasories involving aberrant protein expression or biological activity. The polypeptide and polymuclectide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations creponable for genetic disorders or other tratts to assess biodiversity and to produce other types of data and products dependent on DNA and can opposite amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO can fire.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                   New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia; coronary heart disease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              71.3%; Score 97; DB 22; Length 52; 63.2%; Pred. No. 7.1e-06; tive 4; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 20; SEQ ID No 47452; 103pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Human liver peptide, SEQ ID No 30819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABG52171 standard; Peptide; 292 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 ККНИНКИНККИНИКИ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19 ORHHEITHEIDYHHEITHO 37
                                                                                                                                                                                                                                                        Tang YT
                                                                                                                                   31-MAR-2000; 2000US-0540217.
23-AUG-2000; 2000US-0649167.
                                                                                          30-MAR-2001; 2001WO-US08631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        25-FEB-2003 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 63.2'
Matches 12; Conservative
                                                                                                                                                                                                                                                      Drmanac RT, Liu C,
                                                                                                                                                                                                                                                                                                         WPI; 2001-639362/73.
                                                                                                                                                                                                          (HYSE-) HYSEQ INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  52 AA;
                                                                                                                                                                                                                                                                                                                            N-PSDB; AAS81280
WO200175067-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200157273-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                 biodiversity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Homo sapiens
                                           11-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ABG52171;
```

RESULT 13 ABG52171

ઠ g ö

Gapa

Rank DR;

```
Human; foetal liver; gene expression; single exon nucleic acid probe.
                                                                                                                                                                                                                                                                                                                                                        Peptide #4866 encoded by human foetal liver single exon probe
                                                                                                                                                                                                                                                                                                                     ABB37360 standard; Peptide; 292 AA.
                                                                                                                                                                                                                                                                          3 нинкиннккиникк 20
                                                                                                                                                                                                                                                                                 (MOLE-) MOLECULAR DYNAMICS INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                           2000US-0207456.
2000US-0608408.
2000US-0632366.
                                                                                                                                                                                                                                                                                                                                                                                                                          30-JAN-2001; 2001WO-US00669
                                                                                                                                                                                                                                                                                                                                                                                                                                                              2000US-0234687
                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 66.7
Matches 12; Conservative
                                          Hanzel DK,
                                                      WPI; 2001-496933/54
                                                                                                                                                                                                                                      292 AA;
                                                                                                                                                                                                                                                                                                                                                                                                 WO200157277-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                   30-JUN-2000;
03-AUG-2000;
21-SEP-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                       04-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                              26-MAY-2000;
                                                                                                                                                                                                                                                                                                                                                                                                              09-AUG-2001.
                                                                                                                                                                                                                                                                                                                                               04-FEB-2002
     21-SEP-2000;
27-SEP-2000;
                04-OCT-2000;
                                                                                                                                                                                                                                                                                                                                   ABB37360;
                                                                                                                                                                                                                                      Sequence
                                           Penn SG,
                                                                                                                                                                                                                                                                                                                 ABB37360
                                                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                             ઠે
```

```
Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human fetal liver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Search completed: January 20, 2004, 18:23:19
Job time : 32.0145 secs
                                                                             (MOLE-) MOLECULAR DYNAMICS INC.
                                                                                                                                Chen W,
27-SEP-2000; 2000US-0236359.
04-OCT-2000; 2000GB-0024263.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 66...
Local 12; Conservative
                                                                                                                                Hanzel DK,
                                                                                                                                                                                   WPI; 2001-483447/52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                     SG,
                                                                                                                                        Penn
             88X8X8X8X8X8X8X8X8X8X8X8X8X8X8X8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                       The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived nucleic acid probes for measuring gene expression in a sample derived them becast and Br 474 cells. The method involves contacting the probes with a collection of detectably labelled nucleic acids derived from mRNA of human breast, and then measuring the label derived from mRNA of the microarray. The probes are useful for bound to each probe of the microarray. The probes are useful for encode proteins. They are useful for gene discovery, and for encode proteins. They are useful for gene discovery, and for expression analysis is useful for assessing the toxicity of chemical expression analysis is useful for assessing the toxicity of chemical diversity of probes for measuring gene expression, with far less bias diversity of probes for measuring gene expression, with far less bias than expressed sequence tag microarrays. The method is suitable for rapid production of functional information from genomic sequence. The report of the formation from sequence. The formation from micleic acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                              from human
                                                                                                                                                                                                                                                                                              New spatially-addressable set of single exon nucleic acid probes, useful for measuring gene expression in sample derived from human breast, comprises number of single exon nucleic acid probes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                 Claim 27; SEQ ID NO 15076; 327pp + sequence listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    71.3%; Score 97; DB 22; Length 292; 66.7%; Pred. No. 4e-05;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      probe of the invention.
Note: The sequence data for this patent
                                                                                                                                                                                                      Chen W, Rank
                        ; 2000US-0632366.
; 2000US-0234687.
; 2000US-0236359.
; 2000GB-0024263.
                             03-AUG-2000;
```

```
The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probbes may be used for predicting, measuring and displaying gene expression in samples derived from human fetal liver. The present sequence is a peptide encoded by a single exon nucleic acid probe of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
Claim 27; SEQ ID NO 29995; 639pp + sequence listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 97; DB 22; Length 292;
Pred. No. 4e-05;
2; Mismatches 4; Indel8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 НИНКИННККИНИКИ 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ||| ||| ||| |||:
81 HHPHHNHHHHHHRR 98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       71.3%;
```

```
US-08-203-532F-2
ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 2, Appli
Sequence 4, Appli
Sequence 5, Appli
Sequence 16, Appli
Sequence 6, Appli
Sequence 7, Appli
Sequence 7, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                         January 20, 2004, 18:22:04; Search time 11.3043 Seconds (without alignments) 74.858 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 4,
Sequence 17
Sequence 1,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2
Sequence 2
Sequence 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence Sequence Sequence Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence Sequence Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /cgn2 6/ptodata/1/iaa/5A_COMB.pep:*
/cgn2 6/ptodata/1/iaa/5B_COMB.pep:*
/cgn2 6/ptodata/1/iaa/6A_COMB.pep:*
/cgn2 6/ptodata/1/iaa/6A_COMB.pep:*
/cgn2 6/ptodata/1/iaa/PCTUS_COMB.pep:*
/cgn2 6/ptodata/1/iaa/PCTUS_COMB.pep:*
                       GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-252-991A-30855
US-07-882-292-2
US-08-331-644-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-328-352-4930
US-08-203-232F-2
US-09-078-465-2
US-09-078-465-2
US-09-312-295-4
US-09-312-295-4
US-09-312-295-4
US-09-312-295-4
US-09-312-295-4
US-09-315-032-1
US-09-255-457-1
US-09-15-032-1
US-09-255-571-2
US-09-378-46-2
US-09-275-4
US-08-203-532F-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JS-08-557-006C-43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-599-652-2
US-08-642-846-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                            328717 segs, 42310858 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                            OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                           US-10-018-103A-6
136
1 КИННИННИКИНКК 20
                                                                                                                                                                                                                                                                                                                                                                                                          Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                  BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  75.5
75
75
75
75
75
74
74
                                                                                                                                                                                                                                                                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                                                                          Perfect score:
                                                                                                                                                                                                                                                                                                                                  Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Database
                                                                                                                                                                              Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Result
```

```
RESULT 1
US-09-328-352-4930
; Sequence 4930, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION UNMERR: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SED ID NO 4330
; LENGTH: 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                    Sequence 3, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 12, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 21, Appli
Sequence 31, Appl
Sequence 36, Appl
Sequence 36, Appl
Sequence 36, Appl
Sequence 35, Appl
Sequence 25, Appli
Sequence 25, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 363;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 2, Application US/08203532F
Patent No. 5856121
GENERAL INFORMATION:
APPLICANT: GGTEK!, David H.
TITLE OF INVENTION: Growth Arrest Homeobox Gene
NUMBERSONDENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calfee, Halter, and Griswold
STREET: 800 Superior Avenue
STREET: Gleveland
STRTE: Olio
US-09-264-604-2

US-08-686-28A-3

US-09-456-287-3

US-08-686-528A-2

US-09-461-474-12

PCT-US92-06640-2

US-08-311-023-2

US-08-311-023-2

US-08-461-474-8

US-09-461-474-8

US-09-461-474-8

US-09-461-474-8

US-09-461-474-8

US-09-461-325-160

US-09-126-980-2

US-09-126-980-2

US-09-126-980-2

US-09-126-980-2

US-09-126-980-2

US-09-461-482-2

US-09-126-980-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUMTRY: U.S.A.
ZIP: 44114-268

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/NS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/203,532F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 93; DB 4; ] Pred. No. 0.00013;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                             ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 188 HHTHHHHHHHHE 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  нникниккиникник 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 70.6%;
Matches 12; Conservative
    1664
313
313
313
313
313
313
444
443
765
765
1118
1115
726
726
```

```
Gaps
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 303;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Kaudimann, Joerg
TITLE OF INVENTION: CIF150/hTAFII150 is Necessary for Cell
TITLE OF INVENTION: Cycle Progression
FILE REFERENCE: 1453.002
CURRENT APPLICATION NUMBER: US/09/208,742
CURRENT APPLICATION NUMBER: 198-12-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSEQ for Windows Version 3.0
  Indels
                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY: U.S.A.
ZIP: 44114-2688
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: PREMENT PC-DOS/NS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01882A
                                                                                                                                                                                                               APPLICANT: Gorski, David H.
APPLICANT: Gorski, David H.
APPLICANT: Walsh, Kenneth
TITLE OF INVENTION: Growth Arrest Homeobox Gene
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        63.2%; Score 86; DB 5; 1
57.9%; Pred. No. 0.00075;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3; Mismatches
                                                                                                                                                                                                                                                                                                                         and Griswold
  3; Mismatches
                                                                                                                                                                                Sequence 2, Application PC/TUS9501882A GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 2, Application US/09208742
Patent No. 6174679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 КНИНИНИКНИККИНКК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  :|| ||| ||| ||| ||| 63 ОННЯ СНИНИНИНИ В 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: GOLITICK, MARY E.
REGISTRATION NUMBER: 34829
REFERENCE/DOCKET NUMBER: 223
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 622-8200
TELEFAX: (216) 241-0816
                                          2 КИННКИННККИННКИ 20
                                                                     63 ОНИКСИМИНИНИНИНОО 81
                                                                                                                                                                                                                                                                                                                           ADDRESSEE: Calfee, Halter, STREET: 800 Superior Avenue CITY: Cleveland
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELBFAX: (216) 241-0816
TELBEX: 980499
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     : 303 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11; Conservative
      11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
ORGANISM: human
                                                                                                                                                                                                                                                                                                                                                                                           STATE: Ohio COUNTRY: U.8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQ ID NO 2
            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 86; DB 3; Length 303; Pred. No. 0.00075;
                                                                                                                                                                                                                                                                                                                                                                          Query Match 63.2%; Score 86; DB 2; Length 303; Best Local Similarity 57.9%; Pred. No. 0.00075; Matches 11; Conservative 3; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
APPLICANT: Gorski, David H.
APPLICANT: Walsh, Kenneth
TITLE ON INVENTION: Growth Arrest Homeobox Gene
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA:
FILING DATE: 24-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: GOLICK, MATY E.
REGISTRATION NUMBER: 34829
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEFRAX: (216) 622-8200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE: Calfee, Halter, and Griswold
STREET: 800 Superior Avenue
CITY: Cleveland
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY: U.S.A.
ZIP: 44114-2688
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: PATENT PC-DOS/MS-DOS
SOFTWARE: PATENT Release #1.0; Ve)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/078,465
FILING DATE: 14-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 2, Application US/09078465
Patent No. 6280969
                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 КНННКНННККНННКК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: GOLIACK, MAIY E.
REGISTRATION NUMBER: 34829
REFRENCE/DOCKET NUMBER: 2231
TELEFORMUNICATION INFORMATION:
TELEFHONE: (216) 622-8200
TELEFAX: (216) 241-0816
TELEFAX: (216) 241-0816
TELEX: 980499
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 303 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : 303 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEX: 980499
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                ; TOPOLOGY: linear; MOLECULE TYPE: protein US-08-203-532F-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEFAX:
TELEX: 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE: OF COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-078-465-2
```

ઠે 셤

```
Kroeger, Burkhard
Kuenast, Christoph
ENTION: No. 5523287el thrombin-inhibitory protein from assassin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-255-457-1
; Sequence 1, Application US/08255457
; Patent No. 5780400
; Patent No. 5780401
; APPLICANT: Plaut, Andrew G.
; APPLICANT: Wright, Andrew; TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL BINDING TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS: 3
; CORRESPONDENCE ADDRESS: 5
; STREET: 225 Franklin Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 83, DB 1; Length 368;
Pred. No. 0.0021;
1; Mismatches 5; Indels
                                                                                                                                                                                                                                   COMPUTRY: USA
ZIP: 20036
ZIP: 20036
ZIP: 20036
COMPUTRY READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage COMPUTER: IBM AT-compatible, 80486 processor OPERATING SYSTEM: MS-DOS version 6.0
SOFTWARE: Worderfect version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/211,942
FILING DATE: VERSITION COTK 13/00
CLASSIFICATION: COTK 13/00
CLASSIFICATION AGIK 37/64
PRIOR APPLICATION DATA:
APPLICATION DATA: PCT/EP92/02450
FILING DATE: 27-0CT-1992
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
**TWATH-** 368 mmino acid8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY: U.S.A.

ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Ploppy disk
COMPUTER: IBM PC Compatible
COMPUTER: Patentin PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/255,457
                                                                                                                                                                  STREET: 1101 Connecticut Avenue CITY: Washington STATE: D.C.
                   APPLICANT: Kroeger, Burkhard
APPLICANT: Kuenast, Christoph
TITLE OF INVENTION: No. 5523287
TITLE OF INVENTION: bugs.
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Keil & Weinkauf
Bialojan, Siegfried
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul C.
REGISTRATION NUMBER: 30,162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            176 ОНИНИНРРРИНИНИ 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 64.7%;
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CITY: Boston
STATE: Massachusetts
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-211-942-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61.8%; Score 84; DB 4; Length 1199; 60.9%; Pred. No. 0.0049;
                                             Length 1199;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61.8%; Score 84; DB 4; Length 1199; 60.9%; Pred. No. 0.0049; tive 0; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Kaufmann, Joerg
TITLE OF INVENTION: CIF130 INHIBITS CELL CYCLE PROGRESSION
FILE REPERENCE: 200130.456 / 1513.003
CURRENT APPLICATION NUMBER: US/09/332,295
CURRENT FILING DATE: 1999-06-11
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PastSEQ for Windows Version 3.0
SEQ ID NO EBO ID NOS: 4
LENGTH: 1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 4, Application US/09709979
PARENT NO. 6423822
GENERAL INFORMATION:
APPLICANT: Kaufmann, Joerg
TITLE OF INVENTION: CIFIJO INHIBITS CELL CYCLE PROGRESSION
FILE REPERENCE: 200130.456 / 1513.003
CURRENT APPLICATION NUMBER: US/09/709,979
CURRENT PILING DATE: 2000-11-09
PRIOR PILING DATE: 1999-06-11
                                        Query Match 61.8%; Score 84; DB 3; Best Local Similarity 60.9%; Pred. No. 0.0049; Matches 14; Conservative 0; Mismatches 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 3.0
                                                                                                                                                                                 1146 ніннінініненккккікнікнік 1168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1146 НИНИНИЕНКККККККНКИКНК 1168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1146 НИНИНИЕНКККККККНКНКНК 1168
                                                                                                                                      3 нининин-----Кинини 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 нинкинн-----кининик 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 нинхини-----ккиникиник 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 17, Application US/08211942
Patent No. 5523287
GENERAL INFORMATION:
APPLICANT: Friedrich, Thomas
                                                                                                                                                                                                                                                                         US-09-332-295-4

Sequence 4, Application US/09332295

; Patent No. 6303372

; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 60.9
Matches 14; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 60.9
Matches 14; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; ORGANISM: Homo sapien
US-09-332-295-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; ORGANISM: Homo sapien
US-09-709-979-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQ ID NO 4
LENGTH: 1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-211-942-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-709-979-4
US-09-208-742-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                        ઠે
                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
```

```
FILING DATE
                                                                                                                                                                                                                                                                                                        ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
US-08-677-862-2
                                                                                                         PCT-US95-05772-1
                                                                                                                                                                                                                                                                                                                             STREET:
                                                                                    RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-115-032-1
; Sequence 1, Application US/09115032
; Sequence 1, 5972348
; Patent No. 5972348
; Patent No. 5972348
; APPLICANT: Plaut, Andrew G. APPLICANT: Gilbert-Rochstein, Joanne V. APPLICANT: Wright, Andrew G. TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL BINDING TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL BINDING COMMERCE OF SEQUENCES: 3
CORRESPONDENCE ADDRESS: 3
CORRESPONDENCE ADDRESS: ADDRESSEE: Fish & Richardson STREET: 225 Franklin Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 60;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3; Indels
                                                                                                                                                                                                                                                                     Length 60;
                                                                                                                                                                                                                                                                                                              3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY: DUBLING MASSACHUSELES
COUNTRY: U.S.A.
ZIP: 0210-2804
ZIP: 0210-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/115,032
FLING DATE:
CLASSIFICATION DATA:
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 82; DB 2;
Pred. No. 0.00048;
1; Mismatches 3
                                                                                                                                                                                                                                                                       60.3%; Score 82; DB 1; ilarity 60.0%; Pred. No. 0.00048; Conservative 1; Mismatches 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul C.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
                    00398/090001
                                                                                                                                                                                                                                                                                                                                                                                3 ННИКИНКИНИН 18
                REFERENCE/DOCKET NUMBER: 0039
TELECOMMUNICATION INFORMATION:
TELERAX: (617) 542-5070
TELERAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 60.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 60.00
Best Local Similarity 60.00
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEX: 200154
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                60 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      , MOLECULE TYPE: protein US-09-115-032-1
                                                                                                                                                                                                                    MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 12; Conserval
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     amino acid
                                                                                                                                                                                                   linear
                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                                                                          US-08-255-457-1
                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                  ઠે
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 2, Application US/08677862; Sequence 2, Application US/08677862; Sequence 2, Application US/08677862; Setent No. 5874230; Patent No. 5874230; APPLICANT: SOUGH, HO Yeong APPLICANT: ROTHE, Mike Mike TRAE2-Associated Kinase TITE OF INVENTION: TRAE2-Associated Kinase; NUMBER OF SEQUENCES: CORRESPONDENCE ADDRESS: ADDRESSE: Flehr, Hobbach, Test, Albritton & Herbert STREET 4 Banacradero Center, 34th floor CITY: San Francisco COUNTRY: USA COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 60;
                                                                                                                                                                                                                                                                                    COMPUTEY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 82; DB 5; 1 Pred. No. 0.00048;
Sequence 1, Application PC/TUS9505772
GENERAL INFORMATION:
APPLICANT: Plaut, Andrew G.
APPLICANT: Gilbert-Rothstein, Joanne V.
APPLICANT: Mright, Andrew
TITLE OF INVENTION: HELICOBACTER PYLORI NICKEL
TITLE OF INVENTION: BINDING PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION:
ATTORNEY/ACENT INFORMATION:
NAME: CLASK, Paul C.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00398/090001
TELECOMMUNICATION INFORMATION:
TELEFAX: (617) 542-5070
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11 ННННННННННН 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 ННИКНИНККИНН----КИНН 18
                                                                                                                                                                                                    Fish & Richardson
                                                                                                                                                                                                                          225 Franklin Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                60.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 60.03
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEX: 200154
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: protein
                                                                                                                                                                                                                                              CITY: Boston
STATE: Massachusetts
                                                                                                                                                            NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    inear
```

ä

```
Gaps
                                                                                                                 3;
                                                                    Score 81.5; DB 2; Length 763;
Pred. No. 0.0063;
0; Mismatches 4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match

59.9%; Score 81.5; DB 3; Length 763;
Best Local Similarity 63.2%; Pred. No. 0.0063;
Matches 12; Conservative 0; Mismatches 4; Indels
                                                                                                                                                                                                                                                                                                      Sequence 2, Application US/09434065
; Sequence 2, Application US/09434065
; Patent No. 6107074
; GENERAL INFORMATION:
; APPLICANT: SONG, Ho Yeong
; TITLE OF INVENTION: TRAF2-Associated Kinase
; TITLE OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, 34th floor
; CITY: San Francisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/434,065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brezner, David J
REGISTRATION NUMBER: 24,774
REFERENCE/DOCKET NUMBER: T96-005/A63613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415 781-1989
TELEPRAX: 415 398-3249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER READABLE FORM:
MEDIUM TYPE: PLOSDY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 4, Application US/08789275A
Patent No. 6251664
GENERAL INFORMATION:
APPLICANT: Palleja Estivill Xavier
APPLICANT: Pritchard, Melanie
APPLICANT: Vilaro, Jordi Guimera
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/677,862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 600 HHHHHHHHHHHHH 618
                                                                                                                                                                                                        600 нинсизэннининин 618
                                                                                                                                                              3 НЯН---КИНИКИНИНИ 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 нин---киниккиникнии 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 not relevant
                                                                    Query Match 59.9%;
Best Local Similarity 63.2%;
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 763 amino acide
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
  ; MOLECULE TYPE: peptide US-09-252-571-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CA
USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: US
ZIP: 94111
                                                                                                                                                                                                                                                                                                  US-09-434-065-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-434-065-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 15
US-08-789-275-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE:
                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 59.9%; Score 81.5; DB 2; Length 763; Best Local Similarity 63.2%; Pred. No. 0.0063; Matches 12; Conservative 0; Mismatches 4; Indels 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-252-571-2

Sequence 2, Application US/09252571

Sequence 2, Application US/09252571

Settler No. 5981250

APPLICANT: SONG, HO Yeong

APPLICANT: SONG, HO Yeong

APPLICANT: ROTHE, Mike

TITLE OF INVENTION: TRAP2-Associated Kinase

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSE: Flehr, Hobbach, Test, Albritton & Herbert

STREET: 4 Embarcadero Center, 34th floor

CITY: San Francisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/NS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/252,571
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/677,862
FILING DATE:
                                                                                                                          CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brezner, David J
REGISTRATION NUMBER: 24,774
REDESTRATION NUMBER: 196-005/A63613
TELECOMMUNICATION INFORMATION:
TELEPAX: 415,398-3249
INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS:
LENGTH: 763 aning acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/677,862
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brezner, David J
REGISTRATION NUMBER: 24,774
REFERENCE/DOCKET NUMBER: 796-005/A63613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415 781-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 600 нін GNSS Нічнін нін 618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 нин---киникинин 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEFAX: 415 398-3249
INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS: LENGTH: 763 amino acids TYPE: amino acid STRANDEDNESS: not relevant TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                   not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                         , MOLECULE TYPE: peptide US-08-677-862-2
                                                                                                                                                                                                                                                                                                                                                                                           TYPE: amino acid
STRANDEDNESS: not
                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE:
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
```

8 셤 us-10-018-103a-6.rai

```
TITLE OF INVENTION: Human Gene Sequence of the Down Syndrome Critical Patent No. 6251664

TITLE OF INVENTION: Region of Human Chromosome 21, Coding for A Serine-TITLE OF INVENTION: Threonine Protein Kinase (MNB), Expressed in the TITLE OF INVENTION: Threonine Protein Kinase (MNB), Expressed in the FILE REFERENCE: U 01114-4

CURRENT APPLICATION NUMBER: US/08/789,275A

CURRENT FILING DATE: 1997-01-28

NUMBER OF SEQ ID NOS: 6

SEQ ID NO 4

LENGH+ 763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 59.9%; Score 81.5; DB 3; Length 763; Best Local Similarity 63.2%; Pred. No. 0.0063; Ai Indels 3; Gaps Matches 12; Conservative 0; Mismatches 4; Indels 3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Search completed: January 20, 2004, 18:27:12
Job time : 11.3043 Bec8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                600 HHHGNSSHHHHHHHHHH 618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  з ннн---княнккиникин 18
                                                                                                                                                                                                                                                                                                                                                ; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-08-789-275-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ઠે
```

Sequence 37, p Sequence 40, Sequence 41, p Sequence 45, p Sequence 43, p Sequence 38, p

Sequence

Sequence 1992, Applisequence 16, Applisequence 16, Applisequence 16, Applisequence 16, Applisequence 16, Applisequence 17, Applisequence 476, Applisequence 17, Applisequence 14065, Applisequence 13185, Asquence 13185, Asquence 61, Applisequence 61, Applisequence 61, Applisequence 61, Applisequence 61, Applisequence 62, Applisequence 65, Applisequence 66, Applisequence 6

Sequence 14, App

Scoring table:

Searched:

Database

Perfect score:

Run on:

Sequence:

```
APPLICANT: Mixson, A. James
TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
FILE REFERENCE: 56274
FULS REPERENCE: 56274
CURRENT APPLICATION NUMBER: US/10/131,909A
CURRENT FILING DATE: 2001-11-05
FRIOR PPLICATION NUMBER: US 10/018103
FRIOR PPLICATION NUMBER: PCT/US00/34603
FRIOR PLING DATE: 2000-12-20
FRIOR APPLICATION NUMBER: PCT/US00/34603
FRIOR PLING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin version 3.1
SEQ ID NO 6
LENGTH: 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
2 US-10-315-515-36

2 US-10-315-515-40

2 US-10-315-515-40

2 US-10-315-515-42

2 US-10-315-515-42

2 US-10-315-515-45

2 US-10-315-515-45

2 US-10-315-515-38

2 US-10-215-515-38

2 US-10-215-2198-16

3 US-10-2198-16

3 US-10-2198-16

4 US-10-2198-16

5 US-10-318-3198-3

6 US-10-318-3198-3

6 US-10-318-3198-3

7 US-10-295-386-34005

2 US-10-295-386-34105

2 US-10-295-386-31185

2 US-10-293-386-31185

2 US-10-293-386-31185

2 US-10-293-386-31185

2 US-10-293-386-31185

2 US-10-293-386-31185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 100.0%; Score 136; DB 12; Best Local Similarity 100.0%; Pred. No. 3.6e-08; Matches 20; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      OTHER INFORMATION: Synthetic Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 6, Application US/10111909A; Publication No. US20030165567A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 2
US-10-018-103A-6
; Sequence 6, Application US/10018103A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 ККИННКИННКИННККК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KKEHHKEHKKKEHKKEHHKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                              12
12
12
13
14
15
   US-10-131-909A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JS-10-131-909A-6
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 6, Appli
Sequence 17, Appl
Sequence 17, Appl
Sequence 31882, A
Sequence 37884, A
Sequence 31914, A
Sequence 31311, A
Sequence 34744, A
Sequence 34744, A
Sequence 34744, A
Sequence 34, Appl
Sequence 44, Appl
Sequence 46, Appl
Sequence 35, Seppl
Sequence 36, Appl
Sequence 36, Appl
                                                                                               January 20, 2004, 18:23:25; Search time 22.6087 Seconds (without alignments) 180.887 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                      '(gnz_6')ptodata'1/pubpaa/USO7_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO7_NEW_PUB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO6_NEW_PUB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO6_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO7_NEW_PUB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO8_NEW_PUB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO8_NEW_PUB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO8_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO8_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO8_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO9_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO9_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO9_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO9_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO9_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO9_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO9_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO9_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO8_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO8_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO8_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO8_PUBCOMB.pep:*
'(gnz_6')ptodata'1/pubpaa/USO8_PUBCOMB.pep:*
                  GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2 US-10-131-909A-6
5 US-10-018-103A-6
10S-10-018-103A-17
2 US-10-029-386-33892
US-09-864-761-37944
US-09-864-761-37944
US-09-864-761-33313
US-09-864-761-33727
US-09-864-761-34744
US-09-864-761-34744
US-09-864-761-34744
US-10-315-515-39
2 US-10-315-515-34
2 US-10-315-515-36
2 US-10-315-515-36
                                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                        762491 segs, 204481190 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                           Published_Applications_AA:*
                                                                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                            OM protein - protein search, using sw model
                                                                                                                                                                   US-10-018-103A-6
136
1 ККННИЧНИКИНКИНКИ 20
                                                                                                                                                                                                                             BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                  Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query
Match Length
```

ö

Gарв

o;

Score

```
PRIOR
PRIOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 17, Application US/10131909A

Sequence 17, Application US/20131909A

Publication No. US20030165567A1

GENERAL INFORMATION:

APPLICANT: Mixson, A. James

TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same

TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same

TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same

CURRENT APPLICATION UNMBER: US/10/131,909A

CURRENT FILING DATE: 2001-11-05

PRIOR APPLICATION NUMBER: US 10/018103

PRIOR PILING DATE: 2000-12-20

PRIOR APPLICATION NUMBER: US 60/173576

PRIOR FILING DATE: 1999-12-29

NUMBER OF SEQ ID NOS: 17

SOFTWARE: Patentin version 3.1

LEMOTHUM: 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
publication No. US20030045465A1

GENERAL INFORMATION:

TTILE OF INVENTION: Histidine Copolymer and Methods For Using Same
TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
TITLE OF INVENTION: Histidine Copolymer and Methods For Using Same
CURRENT APPLICATION WINBER: US/10/018,103A
CURRENT FILING DATE: 2001-11-05
FRIOR APPLICATION NUMBER: US 60/173576
FRIOR APPLICATION NUMBER: PCT/US00/34603
FRIOR PELING DATE: 1999-12-29
FRIOR FILING DATE: 2000-12-20
NUMBER OF SEQ ID NOS: 16

SOFTWARE: Patentin version 3.1

LENGTH: 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 110.5; DB 12; Length 17;
Pred. No. 1.2e-05;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No. 1.2e-05;
0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 136; DB 15;
Pred. No. 3.6e-08;
; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
100.0%; Score 136; I
Best Local Similarity 100.0%; Pred. No. 3.6
Matches 20; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 33892, Application US/10029386 Publication No. US20030194704A1 GENERAL INFORMATION: APPLICANT: Penn, Sharron G. APPLICANT: Rank, David R. APPLICANT: Hanzel, David K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURE:
, OTHER INFORMATION: Synthetic Peptide
US-10-131-909A-17
                                                                                                                                                                                                                                                                                                                                                                                                                                            OTHER INFORMATION: Synthetic Peptide US-10-018-103A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 ККНИНКНИККНИКК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          книнкиникиник 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 94.4%;
Matches 17; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-10-029-386-33892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 3
US-10-131-909A-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
GENERAL INFORMATION:

APPLICANT: Penn, Sharton G.
APPLICANT: Penn, Sharton G.
APPLICANT: Penn, Sharton R.
APPLICANT: Hand, David R.
APPLICANT: Hand, David R.
APPLICANT: Hand, David R.
APPLICANT: Hand, David R.
APPLICANT: Great REPRESSION ANDLYSIS BY MICROARRAY
TITLE OF INVENTION: HINDAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR TITLE OF INVENTION: HINDAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID FROBES USEFUL FOR SPETION NUMBER: US 60/180, 312
RRIOR APPLICATION NUMBER: US 60/180, 312
RRIOR PELING DATE: 2000-05-26
RRIOR PELING DATE: 2000-05-26
RRIOR PELING DATE: 2000-05-26
RRIOR PELING DATE: 2000-05-26
RRIOR APPLICATION NUMBER: US 60/236, 359
RRIOR PELING DATE: 2000-05-27
RRIOR PELING DATE: 2000-05-27
RRIOR PELING DATE: 2000-01-00
RRIOR APPLICATION NUMBER: PCT/USO1/00666
RRIOR APPLICATION NUMBER: PCT/USO1/00667
RRIOR PELING DATE: 2001-01-30
RRIOR PELI
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR TITLE OF INVENTION: EXPRESSION ANALYSIS TWO FILE REFERENCE: AEOMICA-X-2 CURRENT APPLICATION NUMBER: US/10/029,386 CURRENT FILING DATE: 2001-112-20 NUMBER OF SEQ ID NOS: 34288 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1 SEQ ID NO 33892 LENGTH: 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 98; DB 12; Length 378;
Pred. No. 0.0032;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURE:
OTHER INFORMATION: MAP TO AP001751.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.
US-10-029-386-33892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DR APPLICATION NUMBER: PCT/USO1/00050

DR APPLICATION NUMBER: PCT/USO1/00669

DR FILING DATE: 2001-01-30

R PILING DATE: 2001-01-30

DR APPLICATION NUMBER: PCT/USO1/00665

DR APPLICATION NUMBER: PCT/USO1/00668

DR APPLICATION NUMBER: PCT/USO1/00663

DR PILING DATE: 2001-01-30

DR PILING DATE: 2001-01-30

DR APPLICATION NUMBER: PCT/USO1/00661

DR APPLICATION NUMBER: PCT/USO1/00661

DR APPLICATION NUMBER: PCT/USO1/00661

DR PILING DATE: 2001-01-30

DR APPLICATION NUMBER: PCT/USO1/00661

DR PILING DATE: 2001-01-30

DR APPLICATION NUMBER: PCT/USO1/00670

DR PILING DATE: 2001-01-30

DR PILING DATE: 2001-01-30

DR PILING DATE: 2001-01-30

DR APPLICATION NUMBER: US 60/234,687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 37882, Application US/09864761 Patent No. US20020048763A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           331 HHQHHNNHHQHHH 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             72.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   з нинкиниккиникни 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE: 2000-09-21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 75.0
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                        TYPE: PRT
ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-864-761-37882
```

```
APPLICANT: Penn, Sharron G.
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David R.
APPLICANT: Hanzel, David R.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USFFUL FOR TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 33014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEATURE INFORMATION: MAP TO ACO05414.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.5
OTHER INFORMATION: EXPRESSED IN HELLOO, SIGNAL = 6.2
OTHER INFORMATION: EXPRESSED IN HELLOO, SIGNAL = 7.1
OTHER INFORMATION: EXPRESSED IN HETAL LIVER, SIGNAL = 6.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.7
OTHER INFORMATION: EXPRESSED IN BOULT LIVER, SIGNAL = 7.7
OTHER INFORMATION: EXPRESSED IN BOULT LIVER, SIGNAL = 7.9
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.9
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 292;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 96; DB 12; Length 29;
Pred. No. 0.00058;
2; Mismatches 4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               OTHER INFORMATION: MAP TO AL138895.4

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.88

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.8

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 37944
LENGTH: 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ch 71.3%; Score 97; DB 9; Similarity 66.7%; Pred. No. 0.0032; 12; Conservative 2; Mismatches 4
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR PLING DATE: 2001-01-30
PRIOR PLING DATE: 2001-01-30
PRIOR PLING DATE: 2001-01-30
PRIOR PLING DATE: 2000-09-21
PRIOR PLING DATE: 2000-09-21
PRIOR PLING DATE: 2000-06-30
PRIOR PLING DATE: 2000-06-30
PRIOR PLING DATE: 2000-06-30
PRIOR PLING DATE: 2000-06-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-10-029-386-30014
; Sequence 30014, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 ННИМНИКИНИКК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 70.6%;
Best Local Similarity 66.7%;
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: PRT
ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 12; Conservat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-864-761-37944
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GEREKAL INFOGRATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Hanzel, David R.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR TITLE OF INVENTION: HUMBER: US/09/864.761
CURRENT PEPLICANTON NUMBER: US/01-05-23
PRIOR PELLOATION NUMBER: US/01-01-03-03
PRIOR APPLICANTON NUMBER: US/01-01-03-03
PRIOR PILING DATE: 2000-02-04
PRIOR PILING DATE: 2000-00-03-03
PRIOR PILING DATE: 2000-00-03-03
PRIOR PILING DATE: 2000-00-03-03
PRIOR PILING DATE: 2000-10-1-03
PRIOR PILING DATE: 2000-10-1-03
PRIOR PILING DATE: 2001-01-03-03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                     CTHER INFORMATION: MAP TO ACC11186.2

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3

OTHER INFORMATION: EXPRESSED IN HELIO, SIGNAL = 1.3

OTHER INFORMATION: EXPRESSED IN HELIO, SIGNAL = 1.3

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5

OTHER INFORMATION: EXPRESSED IN BOULT LIVER, SIGNAL = 1.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 97; DB 9; Length 49; Pred. No. 0.00071; Aismatches 3; Indels
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FOR DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 37944, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 ННИМЕНИКИНИКИНИ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            71.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 70.6
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                           ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-864-761-37882
                                                                                                                                                                                                                   SEQ.ID NO 37882
LENGTH: 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
```

ઠે 셤 ö

ò

```
APPLICANT: Rank, David R.

APPLICANT: Hansel, David R.

APPLICANT: Hansel, David R.

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aconica.*L.

CURRENT APPLICATION WINNER: US/09/864,761

CURRENT PELIANO DATE: 2001-05-23

PRIOR PELICATION WINNER: US 60/180,312

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-24

PRIOR PLICATION WINNER: US 60/236,359

PRIOR PLICATION WINNER: US 60/236,369

PRIOR PLICATION WINNER: US 60/236,469

PRIOR PLICATION WINNER: US 60/236,408

PRIOR PLICATION WINNER: US 60/236,
                                                                                                                                                                                           ô
                                                                                                                                                                                           Gaps
                                                                                                                                                                                           ö
                                                                                                                   Length 82;
                                                                                                                                                                                           4; Indels
, OTHER INFORMATION: SWISSPROT HIT: Q12329, EVALUE 4.90e+00 US-09-864-761-33313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OTHER INFORMATION: MAP TO AL096757.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL =
                                                                                                                   Score 92; DB 9;
Pred. No. 0.0035;
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-864-761-33727

Sequence 33727, Application US/09864761

Facent No. US20020048763A1

GENERAL INFORMATION:
                                                                                                                                                                                                                                                                               з нинкиниккини 18
                                                                                                                                                                                                                                                                                                                                                           74
                                                                                                                   Query Match 67.6%;
Best Local Similarity 75.0%;
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                               нинининининин 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT
ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Penn, Sharron G.
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Hanzel, David K.
APPLICANT: CHON Wensheng B.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aconica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT APPLICATION NUMBER: 2001-05-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.1 ER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.1 ER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.8 ER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.5 ER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.5 ER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.4 ER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.4 ER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.2 ER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.6 ER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 33313
LENGTH: 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CURRENT FILLING DATE: 2001-05-29
PRIOR PELLAGION NUMBER: US 60/207,456
PRIOR PELLING DATE: 2000-05-26
PRIOR PELLING DATE: 2000-05-26
PRIOR PELLING DATE: 2000-05-26
PRIOR PELLING DATE: 2000-05-36
PRIOR PELLING DATE: 2000-10-04
PRIOR PELLING DATE: 2000-10-04
PRIOR PELLING DATE: 2000-09-27
PRIOR PELLING DATE: 2000-10-04
PRIOR PELLING DATE: 2001-01-30
PRIOR PELLOATION NUMBER: PCT/US01/00665
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR PELLING DATE: 2001-01-30
PRIOR PELLOATION NUMBER: PCT/US01/00661
PRIOR PELLING DATE: 2001-01-30
PRIOR PELLING DATE: 2000-09-21
PRIOR PELLING DATE: 2000-09-21
PRIOR PELLING DATE: 2000-09-30
                                                                                                                                                                                                                                                                                                           Sequence 33313, Application US/09864761
Patent No. US20020048763A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               OTHER INFORMATION: MAP TO AC007263.3
                                                                                       TYPE: PRT
ORGANISM: Homo sapiens
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                        ; FEATURE:
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN BORE MARROW, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN BL44, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
US-09-864-761-34744
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        67.6%; Score 92; DB 12; Length 90; 75.0%; Pred. No. 0.0038;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 87;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 44, Application US/10315515
; Publication No. US20030166190A1
; GENERAL INFORMATION:
; APPLICANT: Wright, David A.
; APPLICANT: Wright, David A.
; TITLE OF INVENTION: NUCLEIC ACIDS RELATED TO PLANT
; TITLE OF INVENTION: NETROBLEMENTS
; FILE REPERENCE: 08411-031001
; CURRENT FILING DATE: 2002-12-10
; CURRENT FILING DATE: 2001-12-10
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: PastSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Wight, David A.
APPLICANT: Wight, David A.
APPLICANT: Voytas, Daniel F.
TITLE OF INVENTION: NUCLEIC ACIDS RELATED TO PLANT:
TITLE OF INVENTION: RETROBLEMENTS
FILE REFERENCE: 08411-031001
CURRENT APPLICATION NUMBER: US/10/315,515
CURRENT FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: US 60/339,060
PRIOR FILING DATE: 2001-12-10
PRIOR FILING DATE: 2001-12-10
PRIOR FILING DATE: 2001-12-10
SRINGREN OF SEQ ID NOS: 168
SOFTWARE: FRASESEQ FOR Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 67.6%; Score 92; DB 9;
Best Local Similarity 75.0%; Pred. No. 0.0037;
Matches 12; Conservative 0; Mismatches 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; Sequence 39, Application US/10315515; Publication No. US20030166190A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 ННИХННИХИННИЗНН 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 НИНХНИНККИНИИ 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62 нининининини 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   37 инпинининини 52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12; Conservative
             ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM: Glycine max
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 11
US-10-315-515-39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-10-315-515-39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQ ID NO 39
LENGIH: 90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Accomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT APPLICATION NUMBER: US 60/180,312
PRIOR APPLICATION NUMBER: US 60/207,456
                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                           Query Match 67.6%; Score 92; DB 9; Length 87; Best Local Similarity 75.0%; Pred. No. 0.0037; Matches 12; Conservative 0; Mismatches 4; Indels
GTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.2
GTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
GTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
GTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
US-09-864-761-33727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Annomax Sequence Listing Engine vers. 1.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRIOR AFFLIATION NUMBER: US 09/520,7450
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR PILING DATE: 2000-05-26
PRIOR PILING DATE: 2000-10-04
PRIOR PILING DATE: 2000-10-04
PRIOR PILING DATE: 2000-10-30
PRIOR PILING DATE: 2001-01-30
PRIOR PILING DATE: 2000-09-21
PRIOR PILING DATE: 2000-09-21
PRIOR PILING DATE: 2000-09-30
PRIOR PILING DATE: 2000-09-31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 34744, Application US/09864761 Patent No. US20020048763A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                             3 НИНХНИНККИШКИНИ 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         62 HHHHHHHHHHHHHH 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Annom SEQ ID NO 34744 LENGTH: 87
                                                                                                                                                                                                                                                                                                                                                                                                                                 ઠે
```

```
ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ch 67.6%; Score 92; DB 12; Length 95; 1 Similarity 75.0%; Pred. No. 0.004; 12; Conservative 0; Mismatches 4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 67.6%; Score 92; DB 12; Length 93; Best Local Similarity 75.0%; Pred. No. 0.0039; Matches 12; Conservative 0; Mismatches 4; Indels
                                                                                                                                                           Query Match 67.6%; Score 92; DB 12; Length 90; Best Local Similarity 75.0%; Pred. No. 0.0038; Matches 12; Conservative 0; Mismatches 4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 46, Application US/10315515
; Publication No. US20030166190A1
; GENERAL INFORMATION:
    APPLICANT: Wright, David A.
; APPLICANT: Wright, David A.
; APPLICANT: Wright, David A.
; TITLE OF INVENTION: RETROBLEMENTS
; TITLE OF INVENTION: RETROBLEMENTS
; FILE REFERENCE: 08411-031001
; CURRENT FILING DATE: 2002-12-10
; RUGH APPLICATION NUMBER: US 60/339,060
; RIOR PILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46
; SEQ ID NO 46
; SEQ ID NO 40
; SE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 14
US-10-315-515-35
Sequence 35, Application US/10315515
Sequence 35, Application US/10315515
Publication No. US20030166190A1
GENERAL INFORMATION:
APPLICANT: Wright, David A.
APPLICANT: Woytas, Daniel F.
TITLE OF INVENTION: NUCLEIC ACIDS RELATED TO PLANT
TITLE OF INVENTION: RETROELEMENTS
FILE REPERENCE: 08411-031001
CURRENT PILING DATE: 2002-12-10
RIOR APPLICATION NUMBER: US 60/339,060
RIOR APPLICATION NUMBER: US 60/339,060
RIOR RELIGATION NUMBER: 2001-12-10
RIOR RELIGATION NUMBER: 201-12-10
RIOR REPERENCE: 2001-12-10
SRIOR RELIGATION NUMBER: 201-12-10
SRIOR REPERENCE: 2001-12-10
SRIOR RELIGATION NUMBER: 201-12-10
SRIOR REPERENCE: 2001-12-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TYPE: PRT
ORGANISM: Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3 нинкиниккеникнин 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         43 HENHHERHERHER 58
                                                                                                                                                                                                                                                                                                                                                                       3 НННКНИНККННККНН 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                     37 НИННИНИНИНИН 52
                                           ; ORGANISM: Oryza sativa
US-10-315-515-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ) ORGANISM: Glycine max US-10-315-515-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 12; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-10-315-515-35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 13
US-10-315-515-46
TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .
qa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                 ઠે
```

3 нинкиниккиникин 18 ||| ||| ||| ||| 43 нининининин 58

8 8

```
REBURT 15

OGS-10-315-515-34

Sequence 34, Application US/10315515

Sequence 34, Application US/2030166190A1

Sequence 34, Application No. US20030166190A1

Sequence 34, Application No. US20030166190A1

Sequence 34, Application No. US20030166190A1

TITLE OF UNUSATION:

TITLE OF INVENTION:

RETROELEMENTS

PILE REPERENCE: 0841-031001

CURRENT APPLICATION NUMBER: US 60/339,060

PRIOR PEDITORION NUMBER: US 60/339,060

PRIOR PEDITORION NUMBER: US 60/339,060

PRIOR FILING DATE: 2002-12-10

NUMBER OF SEQ ID NOS: 168

SOFTWARE: PRAT

CURRENT Arabidopsis thaliana

US-10-315-515-34

OHENGTH: 96

CRGANISM: Arabidopsis thaliana

US-10-315-515-34

OHENGTH: 86

OHENGTH: 86

OHENGTH: 96

MISMATCHES 12; CONSERVATIVE 0; MISMATCHES 4; Indels

ON

3 HHHKHHKKHHKKHHKKHH 18

ON

3 HHKKHHKKHHKKHHKKHH 18

ON

43 HHHKHHHKHHHHHHHHHHH 58

ON

43 HHHKHHHHHHHHHHHHHH 58

ON

5 SEATCH COMPLECE: 22.6087 SECS
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

- protein search, using sw model OM protein January 20, 2004, 18:18:04 Run on:

; Search time 10.1449 Seconds (without alignments) 189.590 Million cell updates/sec

US-10-018-103A-6 136 1 KKHHHKHHHKKHHKKK 20 Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283308 segs, 96168682 residues Searched:

283308 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR 76:* Database

1: pirl: * 2: pir2: * 3: pir2: * 4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

1 [ا ع مد				
No.	Score	Match	Length	DB	ID	Description
-	96	70.6	407	2	G84783	probable pectinest
~	92	67.6		~	A54523	-rich
ო	92	67.6	35	-4	KGZOHL	4
4	91	6.99		7	A84696	
S	88	64.7		7	A29653	
9	87	64.0		~	T13389	
7	86	63.2	303	~	A56837	homeotic protein M
80	98	63.2	303	~	B49122	homeobox protein M
σ	98	63.2		~	A48130	growth arrest-spec
10	85	62.5		~	266671	neuron-derived rec
11	82	62.5		~	871930	neuron-derived rec
12	85	62.5			JC2493	neuron derived orp
13	85	62.5			T13804	44
14	84	61.8			D83483	probable metal tra
15	84	61.8	390		A38565	$\overline{}$
16	84	61.8			T46024	cal p
11	83	61.0			T39712	
18	83	61.0		•	T42516	ч
19	83	61.0		~	T23056	
20		60.7			B64421	conserved hypothet
21		60.7	14	7	T16440	hypothetical prote
22		60.7		~	T05153	_
23	82.5	60.7		~	S15008	gene disco protein
24	82	60.3	09	~	C64698	probable histidine
25	82	60.3		~	H72583	hypothetical prote
56	81.5	59.9	9	7	T04219	
27	81.5	59.9		7	JC4898	w
28	81	59.6	391	~	H86187	_
29	81	59.6	2649	7	T51023	

hypothetical prote	eggshell protein -	hypothetical prote	transcription acti	URBS1 protein - sm	hypothetical prote	homeotic protein G	neuroblast prolife	knob-associated hi	hypothetical prote	phospholipase C (E	phospholipase C (E	hypothetical prote	otxl protein - mou	homeodomain protei	HBF-G2 (HFK-2) pro
T30119	A54530	T20270	A56235	S27473	T16435	A55641	A40721	A29454.	T45059	A40879	B40879	T44684	835345	156547	137451
7	N	~	~	~	~	~	~	~	~	7	Н	7	~	~	7
102	149	351	311	950	83	302	474	657	735	1305	1312	306	355	355	469
58.8	58.8	58.8	58.5	58.5	58.1	58.1	58.1	58.1	58.1	58.1	58.1	57.4	57.4	57.4	57.4
80	80	80	79.5	79.5	79	79	79	79	79	79	79	78	78	7.8	78
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

```
probable pectinesterase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Accession: G84783
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Accession: G84783
A;Accession: G84783
A;Accession: G84783
A;Accession: Draininary
A;Residues: 1-407 <270>
A;Cross-references: GB:AE002093; NID:g4415916; PIDN:AAD20147.1; GSPDB:GN00139
C;Genetics
A;Map position: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           70.6%; Score 96; DB 2; Length 407; 68.4%; Pred. No. 0.00046; ive 1; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 68.4
Matches 13; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
```

1 ККННИКННККННКК 19 ઠ 셤

ö Length 140; Query Match 67.6%; Score 92; DB 2; Length 140 Best Local Similarity 75.0%; Pred. No. 0.00049; Matches 12; Conservative 0; Mismatches 4; Indels

ö

Gaps

```
Cigrecters Francis North - Insured Cigrecters (Application of the England Part Cigrecters of Sept. 1999)
Cigrecters Home Septiens (Man)
Ciprete: 11-Aug-1995 #sequence revision 11-Aug-1995 #text_change 24-Sep-1999
Ciprete: 11-Aug-1995 #sequence revision 11-Aug-1995 #text_change 24-Sep-1999
Ciprete: 11-Aug-1995 #sequence revision in Residual North Cipreters (Man)
Ci
                  Risalles, C.; Valenti, P.; Darlamitsou, A.; Henderson, N.; Campbell, L.; Glover, I submitted to the EMBL Data Library, May 1999
A; Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A; Peference number: 217665
A; A; Accession: T13389
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Rosidues: DNA
A; Residues: 1-121 < CAT>
A; Residues: L-121 < CAT>
A; Residues: Experiminary translated from GB/EMBL/DDBJ
A; Residues: BNBL; AL031581; NID:e1320978; FID:e1426292; FIDN:CAA20894.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hypothetical protein 115C2.10 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C;Accession: T13389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 1291;
- malaria parasite (Plasmodium falciparum)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 82;
                                                                                                                                                                                                                                                                                                                                   A;Molecule type: DNA
A;Residues: 1-82 <LEN>
A;Cross-references: GB:M17028; NID:g160339; PID:g160341
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ch
1 Similarity 46.2%; Pred. No. 0.00083;
12; Conservative 4; Mismatches 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 87; DB 2;
Pred. No. 0.012;
3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 КННН-----КННИКИННКИН 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Cross-references: FlyBase:FBgn0020381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ::|||:|| ||| ||| 552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 ккиникиникини 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    64.0%;
61.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Reference number: S52168
A; Accession: S52168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A,Map position: X
A,Introns: 238/3; 1225/1
A,Note: EG:115C2.10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
Matches 12; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             C, Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     probable zinc transporter [imported] - Arabidopsis thaliana probable zinc transporter [imported] - Arabidopsis thaliana (mouse-ear cress) (; Species: Arabidopsis thaliana (mouse-ear cress) (; Date: 02-Peb-2001 #sequence_revision 02-Peb-2001 #text_change 02-Peb-2001 C; Accession: A84696 #sequence_revision 02-Peb-2001 #text_change 02-Peb-2001 C; Accession: A84696 A; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Ortonin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. Wiernen, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Aritle: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana. A; Reference number: A84420; MUID: 20083487; PMID: 10617197
                                                                                                                                                                                                                                                                                                                                                     C,Accession: A22692
R,Ravetch, J.V.; Feder, R.; Pavlovec, A.; Blobel, G.
Nature 312, 616-620, 1984
Ajtile: Primary structure and genomic organization of the histidine-rich protein of the
A;Reference number: A22692; MUID:85061618; PMID:6095114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Molecule type: DNA
A;Realdues: 1-515 cRAN-
Z;Cross-references: GB:X01469; NID:g9997; PIDN:CAA25698.1; PID:g9999
C;Comment: There are two copies of 16-residue repeats, two copies of 17-residue repeats,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Cross-references: GB: AE002093; NID: G3980394; PIDN: AAC95197.1; GSPDB: GN00139
C; Genetics:
                                                                                                                                                                                                                                                  histidine-rich glycoprotein precursor - Plasmodium lophurae
C;Species: Plasmodium lophurae
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A) Introns: 23/3
A) Introns: 23/3
C) Superfamily: plasmodium histidine-rich protein
C) Superfamily: plasmodium tandem repeat
C) Superfamily: plasmodium tandem repeat
C) Superfamily: plasmodian: signal sequence #status predicted <SIG>
E):1-23/Domain: propeptide #status predicted <PRO>
P;24-47/Domain: propeptide #status predicted <PRO>
P;59-74,75-90/Region: 1-residue repeats
P;59-74,75-90/Region: 17-residue repeats
P;1-107,108-123/Region: 10-residue repeats
P;173-301,312-331/Region: 10-residue repeats
P;173-301,312-331/Region: 10-residue repeats
P;173-301,312-331/Region: 10-residue repeats
P;173-301,312-331/Region: 10-residue repeats
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 91; DB 2; Length 385;
Pred. No. 0.0016;
3; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 351;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 92; DB 1; Length 351 Pred. No. 0.0011; 0; Mismatches 4; Indels
                                                                                                                                                                                                                                           - Plasmodium lophurae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             67.6%; Scc. 75.0%; Pred 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                201 HHHHHHNHKHQHQHHHKE 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        нинкнинккиннкк 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   176 ННННННАРННННН 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            66.9%;
Local Similarity 66.7%;
hes 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 ННИКНИККИНИКНИН 18
                                                                  ніннінна Риннін 25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 75.0
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A,Status: preliminary
A,Molecule type: DNA
A,Residues: 1-385 <STO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A,Gene: At2g29410
A,Map position: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                             10
              m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                             셤
              ò
```

..

Gaps

.. 8

ö

Gaps

ö

ß RESULT us-10-018-103a-6.rpr

b

```
neuron-derived receptor NOR-1 - human
    C; Accession: A48130; S31976
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Accession: S66671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 11
S71930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     homeobox protein Mox-2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Dec-1993 #sequence revision 17-Nov-1995 #text_change 17-Nov-2000
C;Accession: B49122; S41779; S29902
Britin: A.F.; Hu, J.; Crosby, J.; Lalley, P.A.; Noden, D.; Nadeau, J.H.; Wright, C.V. Development 116, 1123-1136, 1992
A;Title: Mox-1 and Mox-2 define a novel homeobox gene subfamily and are differentially e
A;Reference number: A49122; MUID:93201999; PMID:1363541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Note: the complete translation is not shown R;Candia, A.F.; Kovalik, J.P.; Wright, C.V.E. R;Candia, A.F.; Kovalik, J.P.; Wright, C.V.E. A.; Maino acid sequence of Mox-2 and comparison to its Xenopus and rat homologs. A;Reference number: S41469; MUID:94232829; PMID:7909944
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        growth arrest-specific homeobox protein Cvx - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Cross-references: EMBL:Z16406; NID:g57949; PIDN:CAA78899.1; PID:g57950
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Cross-references: EMBL: Z16406; NID: 957949; PIDN: CAA78899.1; PID: 957950
                                                                    A:Residues: 1-7,'R',9-110,'R',112-122,'Q',124-157,'V',159-303 <GR2>
A:Cross-references: EMBL:X82629
C:Genetics:
A:Genetics:
A:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C;Superfamily: unassigned homeobox proteins; homeobox homology C;Keywords: DNA binding; homeobox; nucleus; transcription regulation F;187-243/Domain: homeobox homology <HOX>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 2; Lens. . 0.0044; 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 303;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 303;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Status: preliminary; nucleic acid sequence not shown Molecule type: mRNA
Residues: 1-303 <CAN2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Status: preliminary; nucleic acid sequence not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 86; DB 2;
Pred. No. 0.0044;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        submitted to the EMBL Data Library, October 1992 A; Reference number: S29902
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 86;
Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 КНИНКИНИККНИКК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2 КНИНИНИКИНИКК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        :|| ||| ||| ||| :: OHHRGHHHHHHHHHQQ 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 63.2%;
Best Local Similarity 57.9%;
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Cross-references: EMBL: 216406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Experimental source: C57BL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: mRNA
A; Residues: 186-246 < CAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: mRNA
A; Residues: 1-303 < CAN1>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Status: preliminary
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Accession: B49122 ·
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Accession: S41779
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Accession: S29902
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

```
A;Note: sequence extracted from NCBI backbone (NCBIN:132842, NCBIP:132843)
R;Gorski, D.H.; LePage, D.F.; Patel, C.V.; Copeland, N.G.; Jenkins, N.A.; Walsh, K.
submitted to the EMBL Data Library, October 1992
A;Description: Cvx: A diverged homeodomain gene which is rapidly down-regulated followi
A;Reference number: S31976
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C;Accession: S66671
R;Petropoulos, I.; Part, D.; Ochoa, A.; Zakin, M.M.; Lamas, E.
FEBS Lett. 372, 273-278, 1995
A;Title: NOR-2 (neuron-derived orphan receptor), a brain zinc finger protein, is highly A;Reference number: S66671; MUID:96000221; PMID:7556683
Rigorski, D.H.; LePage, D.F.; Patel, C.V.; Copeland, N.G.; Jenkins, N.A.; Walsh, K. Mol. Cell. Biol. 13, 3722-373, 1993
AA; Title: Molecular cloning of a diverged homeobox gene that is rapidly down-regulated A; Faterence number: A48130; MUID:93268321; PMID:8098844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A,Gene: nor.2
C;Superfamily: unassigned erbA-related proteins; erbA transforming protein homology
C;Keywords: DNA binding; zinc finger
E;292-430/Domain: erbA transforming protein homology #status atypical <ERBA>
E;292-379/Region: DNA binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Cross-references: EMBL:X86003; NID:g1065885; PIDN:CAA59993.1; PID:g1065886
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        neuron-derived receptor NOR-2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Oct_1996 #sequence_revision 13-Mar-1997 #text_change 20-Sep-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                       A;Cross-references: GB:Z17223; NID:g57951; PIDN:CAA78931.1; PID:g57952 A;Experimental source: aorta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C;Superfamily: unassigned homeobox proteins; homeobox homology C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ..
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 303;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 430;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 86; DB 2; Length 303
Pred. No. 0.0044;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 85; DB 2; Length 43C
Pred. No. 0.0078;
2; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 63.2%; Score 86; DB Best Local Similarity 57.9%; Pred. No. 0.00 Matches 11; Conservative 3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            F;187-243/Domain: homeobox homology <HOX>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 F;294-314/Region: zinc finger CCCC motif
F;312-316/Region: P box
F;313-34/Region: zinc finger CCCC motif
F;313-335/Region: D box
F;312-378/Region: A box
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 книнхнинхкиникк 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 нининикиникк 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    62.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Accession: S31976
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-182, S',184-303 <C
A;Cross-references: EMBL:217223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-430 <PET>
                                                                                                                                                           A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-303 <GOR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 66.7
Matches 12; Conservative
```

```
probable metal transporter PA1297 [imported] - Pseudomonas aeruginosa (strain PAO1)

Sispecies: Pseudomonas aeruginosa
C;Species: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 06-Jan-2003
C;Accession: D83483
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; B. R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Coulter, S.D.; Warrener, P.; Hickey, M.J.; B. J. Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ArTitle: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: D83483
A;Acture: preliminary
A;Acceule type: DNA
A;Residues: 1-326 <STO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Objection (Pc) protein - fruit fly (Drosophila melanogaster)
C.Species: Drosophila melanogaster
C.Species: Drosophila melanogaster
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C.Accession: A3856
R.Parc, R.; Hogness, D.S.
Proc. Natl. Acad. Sci. U.S.A. 88, 263-267, 1991
Proc. Natl. Acad. Sci. U.S.A. 88, 263-267, 1991
A.Reference number: A38565; MuID:91095442; PMID:1898775
A;Reference number: A38565; MuID:91095442; PMID:1898775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Cross-references: GB:AE004559; GB:AE004091; NID:g9947228; PIDN:AAG04686.1; GSPDB:GN00
A;Experimental source: strain PAO1
                                                                                                                                                                                                                                                                                                                                                                                                                          Extrement J.E., Lai, Z.C.; Rubin, G.M.
Development 121, 2835-2845, 1995
A;Title: Shortsighted acts in the decapentaplegic pathway in Drosophila eye development A;Reference number: Z17767; MUID:96038094; PMID:7555710
A;Accession: T13804
A;Accession: T13804
A;Accession: T13804
A;Accession: T13804
A;Accession: Preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                       shs protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C;Accession: T13804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Residues: 1-1212 <TRE>
A;Cross-references: EMBL:L42512; NID:g833709; PID:g833710; PIDN:AAC41608.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 326;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 1212;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
61.8%; Score 84; DB 2;
Best Local Similarity 73.3%; Pred. No. 0.0078;
Matches 11; Conservative 0; Mismatches 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 85; DB 2;
Pred. No. 0.02;
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A,Gene: shs
A,Cross-references: FlyBase:FBgn0010460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C,Genetics:
A,Gene: PA1297
C,Superfamily: zinc transporter ZnT-2
     нининин--нининиоо 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  301 НЕНОРНИНЕННИНО 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 ННИКНИНКИНКИНКИ 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          164 HHHHHHHHHHHHHH 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 64.7%;
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 нинхиникин 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C, Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ઠ
                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A,Status: translated from GB/EMBL/DDBJ
A,Molecule type: mRNA
A,ERSIGNES: 318-45 < RBS-
A,Cross-references: GB:LL9343; NID:g450942; PIDN:AAB46395.1; PID:g565379
A,Experimental source: hippocampus
A,Experimental source: hippocampus
C,Superfamily: probable hormone receptor NIO, nuclear; erbA transforming protein homolog
C,Reywords: DNA binding; receptor; zinc finger
F,292-547/Domain: erbA transforming protein homology <RRBA>
F,292-547/Domain: DNA binding #status predicted
F,312-316/Region: D box
F,311-335/Region: A box
F,372-378/Region: A box
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: mRNA
A; Residues: 1-628 < OMES
A; Residues: 1-628 < OMES
A; Residues: 1-628 < OMES
A; Cross=recences: DDBJ:D38530; NID:g1483194; PIDN:BAA07535.1; PID:g643600
A; Cross=recences: DDBJ:D38530; NID:g1483194; PIDN:BAA07535.1; PID:g643600
A; Cross=recences: Description of the control of the co
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C;Species: Rattus norregicus (Norway rat)
C;Accession: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 20-Jun-2000
C;Accession: JC2493; 148179
R;Ohkura, N.; Hijikuro, M.; Yamamoto, A.; Miki, K.
Biochem. Biophys Res. Commun. 205, 1959-1965, 1994
A;Title: Molecular cloning of a novel thyroid/steroid receptor superfamily gene from cul A;Reference number: JC2493; MUID:95110348; PMID:7811288
C;Species: Homo sapiens (man)
C;Date: 04-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 20-Feb-1998
C;Date: 04-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 20-Feb-1998
C;Accession: S71930
R;Ohkura, N.; Ito, M.; Tsukada, T.; Sasaki, K.; Yamaguchi, K.; Miki, K.
Biochim. Biophys. Acta 1308, 205-214, 1996
A;Title: Structure, mapping and expression of a human NOR-1 gene, the third member of th A;Reference number: S71930; MUID:96404972; PMID:8809112
A;Accession: S71930
A;Molecule type: mRNA
A;Residues: 1-625 <OHK>
A;Cross-references: EMBL:D78579
A;Note: DNA was also sequenced
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A Map position: 99

(S. Superfamally: probable hormone receptor N10, nuclear; erba transforming protein homolog (S. Superfamally: probable hormone receptor; zinc finger
(S. Superfamally: probable stransforming protein homology < BRBA>
(S. Superfamally: probable transforming protein homology < BRBA>
(S. Superfamain: erba transforming protein homology < BRBA>
(S. Superfamain: DAA binding #status predicted < DNA>
(S. Superfamain: DAA binding #status predicted < DNA>
(S. Superfamain: DAA binding #status predicted < DNA>
(S. Superfamain: DAA box
(S. Superfamain: 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ñ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62.5%; Score 85; DB 2; Length 628; ilarity 66.7%; Pred. No. 0.011; Conservative 2; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 625;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 2;
0.011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 85; DB;
Pred. No. 0.01.
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             neuron derived orphan receptor-1 - rat
N;Alternate names: NOR-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 95 ніннінін--іннінно 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          з ненкеннккеникк 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            62.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 66.7
Matches 12, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 Similarity
12; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A, Accession: JC2493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local S
Matches 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           С
```

3 НИНКНИНККИННКК 20

ઠે

Search completed: January 20, 2004, 18:24:44 Job time : 10.1449 secs

			-			
				•		
		·				
	- · ·			•		
			•		· V	
. *	3		•.			
		·	v			
	ω.		X			
•			×			
						•
					*	
4 .						
				•		
	go sa gan ea coma. Se sa coma esta esta esta esta esta esta esta est			÷		
				•		
			*			
	*					•
).):
	a					
	*					
	*					
	*					
	*					
	*					
	*					

```
GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
```

д.	
mode	
3	
using	
search,	
protein	
protein	
ĕ	

Run on:

January 20, 2004, 18:06:29; Search time 6.6667 Seconds (without alignments) 141.080 Million cell updates/sec

US-10-018-103A-6

136 1 ККННИКННККННКК 20 Perfect score:

Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_41:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

STRAMADIES

					SUMMARIES		
Result		Query					
No.	Score	Match	Match Length	DB	QI	Description	
п	92	7	(1)	ี้ ฯ	HRPX PLALO	pla	
~	92	67.6	428	ч	FXB2 MOUSE		
m	88	64.7	82	Н	HRP3_PLAFS	P14586 plasmodium	
4	87	4	465	Н	HNF6 HUMAN	Q9ubc0 homo sapien	
ß	87	64.0	465	ч	HNF6 MOUSE	O08755 mus musculu	
9	87	. 64.0	465	٦	HNF6_RAT	P70512 rattus norv	
7	86.5	63.6	485	٦	ONC2_HUMAN	homo	
60	98	63.2	303	-	MOX2 HUMAN	homo	
σ	98	63.2	303	Н	MOX2 MOUSE	mus m	
10	98	63.2	303	Н	MOX2_RAT	P39020 rattus norv	
11	82	62.5	430	Н	NOR2_RAT	rattu	
12	82	62.5	626	ч	NR43_HUMAN	Q92570 homo sapien	
13	82	62.5	628	ч	NR43_RAT	rattu	
	82	62.5	1211	Н	BUN2_DROME		
15	84	61.8	390	-	PC DROME		
16	83	61.0	420	Н	YBE1_SCHPO		
17	83	61.0	469	Н	A2AC_DIDMA	_	
18	82.5	60.7	143	Н	Y970 METJA	0	
19	82.5	60.7	147	Н	YVS9_CAEEL		
20	82.5	60.7	568	Н	DISC_DROME	P23792 drosophila	
21	•	60.7	1321	Н	PGCN_HUMAN		
22	82	60.3	59	Н	HPN HELPY	Q48251 helicobacte	
23	82	60.3	558	~	CBX4 HUMAN	homo	
24	82	60.3	2004	-	CHDB HUMAN	3 homo	
25	•	-:	763	-	DYRA HUMAN	homo	
56	81.5	-:	763	ч		mus m	
27	81.5	٠.	763	-	DYRA_RAT		
28	80		149	٦	EGGS_SCHMA		
29	80		351	Н	CAV2_CAEEL	Q18879 caenorhabdi	
30	80		977	ч	DLP3_RAT	8	
31	79.5	58.5	950	7	URB1_USTMA	P40349 ustilago ma	
32	79	58.1	437	Н			
33	79	58.1	441	Н	PO33_BRARE	Q90436 brachydanio	

	34		58.1	442	-	VATC DROME		Q9v7n5	drosophila	
	35		58.1	474	-	ANA DROME		026307	drosophila	
	36		58.1	620	н	SNF1 CANAL		P52497	candida alb	
	37		58.1	657	-	KNOB PLAF		P06719	plasmodium	
	38		58.1	1312	-	PIP1 DROME		P25455	drosophila	
	39		57.4	355	-	OTX1 MOUSE		P80205	mus musculu	
	40		57.4	355	-	OTX1 RAT		063410	rattus norv	
	41		57.4	469	٦	PXGA HUMAN	_	P55316	homo sapien	
	42		57.4	477	-	FXGB HUMAN	_	P55315	homo sapien	
	43		57.4	480	Н	FXGB_RAT		000939	rattus norv	
	44	78	57.4	481	-	FXGB MOUSE		Q60987	mus musculu	
	45	77	9.99	307	-	YK55_YEAST		P36155	saccharomyc	
						ALIGNMENTS	INTS			
RESU	RESULT 1									
HRPX :	HRPX PLALO		Ê	COCCUTA			**			
2 2	P04929:	3	100	SIANDAKU;	••	ral;	. WH TCC			
H	13-AUG-1987	1987	(Rel.	05,	Created)	ed)				
Į,	13-AUG-1987	1987	_	05,	ast	Last sequence update)	odate)			
占	15-JUL-1999	1999	(Rel.		ast	Last annotation update)	update)			
			-							

Histidine-rich glycoprotein precursor.

Plasmodium lophurae. Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

NCBI_TaxID=5853;

MEDLINE=85061618; PubMed=6095114;

Ravetch J.V., Feder R., Pavlovec A., Blobel G.;

Ravetch J.V., Feder R., Pavlovec A., Blobel G.;

Ravetch J.V., Feder R., Pavlovec A., Blobel G.;

Ravetch J.V., Feder R.,

Portein of the histidine-rich

protein of the malaria parasite Plasmodium lophurae.";

Nature 312:616-620(1984).

-! MSCELLANBOUS: IN THE INTRAERYTHROCYTIC STAGES OF DEVELOPMENT OF

P.LOPHURAE IN DUCKS, THERE IS A SYNTHESIS OF A MAJOR PROTEIN THAT

ACCUMULATES TO COMPRISE AT LEAST 50% OF THE CELLULAR MASS: THE

HISTIDINE RICH PROTEIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).

Malaria; Repeat; Glycoprotein; Signal. EMBL; X01469; CAA25698.1; -. PIR; A22692; KGZQHL. SIGNAL

HISTIDINE-RICH GLYCOPROTEIN. N-LINKED (GLCNAC. . .) (PROB 2 X 16 AA TANDEM REPEATS. 2 X 17 AA TANDEM REPEATS 23 47 40 90 90 74 1123 1123 1153 1153 1153 48 40 40 53 75 91 CARBOHYD DOMAIN REPEAT CHAIN RESERVATION OF THE PROPERTY OF

(PROBABLE) .

351 18 X 10 AA TANDEM REPEATS. 44032 MW; D19A48D47D890453 CRC64; 2 X 15 AA TANDEM REPEATS. REPEAT DOMAIN REPEAT REPEAT DOMAIN REPEAT REPEAT

351 AA;

SEQUENCE

Gaps ö Score 92; DB 1; Length 351; Pred. No. 0.00097; 0; Mismatches 4; Indels 67.6%; 12; Conservative Query Match Best Local Similarity Matches

ö

3 нинхнинкини 18

ð

```
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                             Malaria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HNF6 HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                              8
                  g
ò
                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                         Kaestner K.H., Schuetz G., Monaghan A.P.; "Expression of the winged helix genes fkh-4 and fkh-5 defines domains
                                                                                                                                                                                                                                                                                                             Kaestner K.H., Lee K.H., Schloendorff J., Hiemisch H.,
Monaghan A.P., Schuetz G.,
"Six members of the mouse forkhead gene family are developmentally
                                                                                                                                            Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 92; DB 1; Length 428;
Pred. No. 0.0012;
0; Mismatches 4; Indel8
                                                                                                                                                                                                                                                                                                                                                  regulated...

Proc. Natl. Acad. Sci. U.S.A. 90:7628-7631(1993).

-! SUBCELLULAR LOCATION: Nuclear.

-! DEVELOPMENTAL STREES: EXPRESSED DURING EMBRYOGENESIS.

-! SIMILARITY: Contains 1 fork-head domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DNA-binding; Nuclear protein; Transcription regulation.
DNA BIND 12 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB8A8EFD1E94AB10 CRC64;
                                                                      01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-NOY-2000 (Rel. 39, Last annotation update)
Forkhead box protein B2 (Transcription factor FKH-4).
                                                             428 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FORK-HEAD.
POLY-HIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLY-ALA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLY-ALA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HSSP, 063245, 24FH.

TRANSFAC; T02442;

MGD: MGI:1347669; Poxb2.

MGD: MGI:1347669; Poxb2.

PÉAM; PP00250; FORK head; 1.

PRINTS; PR00053; FORKHEAD.

PRODOM: PR00339; FH; T.

PROSITE; PS00659; FORK HEAD 1; 1.

PROSITE; PS00659; FORK HEAD 1; 1.

PROSITE; PS00659; FORK HEAD 1; 1.
                                                            PRT;
                                                                                                                                                                                                    SEQUENCE FROM N.A. MEDLINE=97014266; PubMed=8861101;
                                                                                                                                                                                                                                                                                                        MEDLINE=93361500; PubMed=7689224;
                                                                                                                                                                                                                                                nervous system.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; X92591; CAA63335.1; -.
EMBL; X71942; CAAS0744.1; -..
PIR; D47746; D47746.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                45170 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       67.6%;
75.0%;
      нинининарининин 191
                                                                                                                                                                                                                                                                          [2]
SEQUENCE OF 4-114 FROM N.A.
                                                                                                                                                                                                                                                           Mech. Dev. 55:221-230(1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    139
156
163
1217
249
321
396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
Matches 12; Conserv
                                                                                                                                                                                                                                                  in the central
                                                            FXB2_MOUSE
Q64733;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NIMMO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@ib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. Furtha H., Horikawa Y., Wang Y.-Q., Bell G.I.; Furnta H., Horikawa Y., Wang Y.-Q., Bell G.I.; Isolation and characterization of the human hepatocyte nuclear factor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9UBCO; Q99744; Q9UMR6; 16-0CT-2001 (Rel. 40, Created) 16-0CT-2001 (Rel. 40, Last sequence update) 28-PEB-2003 (Rel. 41, Last annotation update) 28-PEB-2003 (Rel. 41, Last annotation update) ONECUTI ON HNP6A OR HNP6.
                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-8720765; PubMed=3038111;
MEDLINE-8720765; PubMed=3038111;
Lenstra R., D'Auriol L., Andrieu B., le Bras J., Galibert F.;
Lenstra R., D'Auriol L., Andrieu B., le Bras J., Galibert F.;
Lenstra R., D'Auriol L., Andrieu B., le Bras J., Galibert F.;
Lenoing and sequencing of Plasmodium falciparum DNA fragments
containing repetitive regions potentially coding for histidine-rich
proteins: identification of two overlapping reading frames.";
Biochem. Biophys. Res. Commun. 146:368-377(1987).
-i- MISCELLANEOUS: THIS PROTEIN IS CODED ON THE REVERSE STRAND OF AN
ASPARTIC ACID-RICH PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
Furnta H., Wang Y.-Q., Bell G.I.;
"The sequence of human mRNA for the hepatocyte nuclear factor-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8;
                                                                                                                                                                                                                                                                                                                          Plasmodium falciparum (isolate fcm17 / Senegal).
Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=5845;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        64.7%; Score 88; DB 1; Length 82;
46.2%; Pred. No. 0.00067;
ive 4; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6 gene.";
Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       82 AA; 10578 MW; 7CE2EA69F2FC1E8C CRC64;
                                                                                                                                                                                                                                01-JAN-1990 (Rel. 13, Created)
01-JAN-1990 (Rel. 13, Last sequence update)
01-MAY-1991 (Rel. 18, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         465 AA
                                                                                                                                                                                    82 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 КННН-----КННККННККННКК 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                        147 НННННААННННН 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M17028; AAA29619.1; -.
3 нинкиниккиникнин 18
                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                       Histidine-rich protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIR; A29653; A29653.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HNF6 HUMAN
                                                                                                                                                                                         HRP3 PLAFS
P14586;
```

m

```
DOMAIN
                                                                                                 ઠે
                              g
                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profilt institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                SEQUENCE OF 1-168 FROM N.A., AND VARIANT ALA-75.
MEDLINE-99420592; PubMed=10491763;
Mocbller A.M., Ek J., Durviaux S.M., Urhammer S.A., Clausen J.O.,
Elberg H., Hansen T., Rousseau G.G., Lemaigre F.P., Pedersen O.;
Hapacocyte nuclear factor-6: associations between genetic variability
and type II diabetes and between genetic variability and estimates of
insulin secretion.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MIM; 604164; -.
GO; 00005634; C:nucleus; NAS.
GO; 00:0005634; F:transcriptional activator activity; NAS.
GO; 00:0016563; F:transcriptional activator activity; NAS.
GO; 00:0006355; P:regulation of transcription, DNA-dependent; NAS.
InterPro; IPR0013108; Cut_homeo.
InterPro; IPR001350; Homeo CUT.
InterPro; IPR001356; Homeo CUT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS00027; HOMEOBOX_1; FALSE_NEG.
PROSITE; PS50071; HOMEOBOX_2; 1.
Transcription regulation; Homeobox; DNA-binding; Nuclear protein;
                                                                                                                                  Samadani U., Costa R.H.; "Yeast one-hybid cloning of the partial human cDNA for hepatocyte nuclear factor 6.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 87; DB 1; Length 465; Pred. No. 0.0042; Aismatches 4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CUT.

HOMBOBOX.

POLY-HIS.

POLY-SER.

P -> A.

FTIG=VAR. 010729.

A -> T IN REF. 3 AND 4).

S -> N (IN REF. 4).

Q -> H (IN REF. 4).

R -> K (IN REF. 4).

K -> Q (IN REF. 4).
Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                EMBL; AP035581; AAD02033.1; -.
EMBL; AF035580; AAD02033.1; JOINED.
                                                                                                 Diabetologia 42:1011-1016(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ProDom; PD000010; Homeobox; 1.
SMART; SM00389; HOX; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      51023 MW;
                                                                                                                        SEQUENCE OF 174-465 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           64.0%;
                                                                                                                                                                                                                                                                                                                                                                                                        EMBL, U96173, AAD00826.1; -. EMBL, X17739, CABS0769.1; -. EMBL, U77975, AAB61705.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF02376; CUT; 1.
Pfam; PF00046; homeobox; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                        TRANSFAC; T03286; -. Genew; HGNC:8138; ONECUT1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Activator; Polymorphism.
DAA_BIND 283 369
DNA_BIND 283 369
DNA_BIND 285 444
DOWAIN 69 72
DOWAIN 124 138
DOWAIN 455 460
VARIANT 75 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 64.7
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     465 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
```

ö

Gaps

ö

2; Mismatches

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Dev. Biol. 192:228-246(1997).
--- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS THE CONSENSUS SEQUENCE
5'-DHWATTGAYTWWD-3' ON A VARIETY OF GENE PROWOTERS SUCH AS THOSE
OF HNF3B AND TTR. IMPORTANT FOR LIVER GENES TRANSCRIPTION (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 R TRANSFAC; T03256; -.

R RGD; MGI:1196423; Onecut1.

RGO; GO:00031015; P:cell differentiation; IMP.

GO; GO:0001492; P:cell differentiation; IMP.

GO; GO:0007492; P:cell differentiation; IMP.

R GO; GO:0007492; P:cell differentiation; IMP.

R GO; GO:0007492; P:cell differentiation; IMP.

R InterPro; IPR001105; Cut_homeo.

R InterPro; IPR001356; Homeo. Cut.

R Pfam; PF00216; Cut; 1.

R Pfam; PF002046; homeobox; 1.

R Probom; PD00010; Homeobox; 1.

R PROBIT: SM00189; HOX: HOMEOBOX 1; FALSE_NEG.

R PROSITE; PS000027; HOMEOBOX 2; 1.

R PROSITE; PS000027; HOMEOBOX 2; 1.

R PROSITE; PS000027; HOMEOBOX 2; 1.
                                                                                                                                                                                                                                                                  16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Hepatocyte nuclear factor 6 (HNF-6) (One cut domain family member 1).
ONECUTI OR HNF6A OR HNF6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=99104231; PubMed=9441664;
Rausa F., Samadani U., Ye H., Lim L., Fletcher C.F., Jenkins N.A.,
Copeland N.G., Costa R.H.;
"The cut-homeodomain transcriptional activator HNF-6 is coexpressed
with its target gene HNF-3 beta in the developing murine liver and
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 87; DB 1; Length 465; Pred. No. 0.0042;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A770D27DD5AAC896 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -!- SUBGNIT: BINDS DNA AS A MONOMER (BY SIMILARITY).
-!- SUBCELLULAR LOCATION: Nuclear.
-!- SIMILARITY: COntains 1 CTT domain.
-!- SIMILARITY: BELONGS TO THE CUT HOMEOBOX FAMILY.
                                                                                                                                                                                                               465 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HOMEOBOX.
POLY-HIS.
POLY-HIS.
                                                   124 НЕННИННИННИРИНОВ 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           50952 MW;
никиникиники 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        64.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; U95945; AAB53863.1; -.
                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              460
                                                                                                                                                                                                                                                                                                                                                                                                                           Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     283
385
69
124
450
465 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE=Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      pancreas.";
                                                                                                                                                                                                               MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DNA_BIND
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA_BIND
```

```
DOMAIN
                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                             ONC2 HUMAN
                                                                                                                                                                                                                                                                                                                    Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
ઠ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce) or send an email to license@lsb-sib.ch).
  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      liver gene expression.";
Mol. Cell. Biol. 16:6273-6284(1996).
-i- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS THE CONSENSUS SEQUENCE
-i- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS THE CONSENSUS SEQUENCE
-i- TRANSCRIPTION AND VARIETY OF GENE PROMOTERS SUCH AS THOSE
OF HIRSB AND TTR. IMPORTANT FOR LIVER GENES TRANSCRIPTION. THE
AFPINITY OF HNF-6ALPHA AND HNF-6BETA FOR DNA DIFFERS DEPENDING ON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ISOUG-P70512-2; Sequence=VSP 002312; TISSUE SPECIFICITY: EXPRESSED IN LIVER, BRAIN, SPLEEN AND TESTIS. SIMILARITY: Contains 1 CUT domain. SIMILARITY: BELONGS TO THE CUT HOMEOBOX FAMILY.
                                                                                                                                           P10512; 080755;
16-0CT-2001 [Rel. 40, Created)
16-0CT-2001 [Rel. 40, Last sequence update)
18-0CT-2003 [Rel. 41, Last annotation update)
28-FEB-2003 [Rel. 41, Last annotation update)
18-FEB-2003 [Rel. 41, Last annotation update)
ONECUTI OR HNF6A OR HNF6.
Rattus norvegicus (Rat).
                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                           "Hepatocyte nuclear factor 6, a transcription factor that contains a novel type of homeodomain and a single cut domain."; proc. Natl. Acad. Sci. U.S.A. 93:9460-9464(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHARACTERIZATION.
MEDLINE=97042457; Bamadani U., Costa R.H.;
Samadani U., Costa R.H.;
"The transcriptional activator hepatocyte nuclear factor 6 regulates
  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Lannoy V.J., Buerglin T.R., Rousseau G.G., Lemaigre F.P.; "Isoforms of hepatocyte nuclear factor-6 differ in DNA-binding "Isoforms of hepatocyte nuclear factor-6 differ in DNA-binding Corperties, contain a bifunctional homeodomain, and define the new ONECUT class of homeodomain proteins."; J. Biol. Chem. 273:13552-13562(1998).
                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A. (ISOFORM ALPHA).
STRAIN=Sprague-Dawley; TISSUB=Liver;
STRAIN=Sprague-Dawley; TISSUB=Liver;
BEDLINE=95182488; PubMed=8790352;
Lemaigre F.P., Durviaux S.M., Truong O., Lannoy V.J., Heuan J.J.,
Rousseau G.G.;
  ;
0
    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Event=Alternative splicing; Named isoforms=2;
     4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IsoId=P70512-1; Sequence=Displayed;
     Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      THE TARGET SEQUENCE.
SUBUNIT: Binds DNA as a monomer.
SUBCELLULAR LOCATION: Nuclear.
ALTERNATIVE PRODUCTS:
                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A. (ISOFORM BETA).
                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN=Wistar;
MEDLINE=98256275; PubMed=9593691;
       ..
7
                                                           124 нининининини 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; X96553; CAA65389.1; -.
EMBL; Y14933; CAA75150.1; -.
TRANSFAC; T03257; -.
                                 4 ННИННИКННИКК 20
        Conservative
                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Name=Alpha;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Name=Beta;
        11;
                                                                                                                                    HNF6 RAT
          Matches
                                                                                                        RESULT 6
                                                                                                                       HNF6 RAT
                                                                                                                                                ð
                                                                g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ..., GO:0003702; F:RNA polymerase II transcription factor acti. . .; TAS. GO:0003397; P:histogenesis and organogenesis; TAS. InterPro; IPR007108; Cut_homeo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Jacquemin P., Lannoy V., Rouseau G.G., Lemaigre F.P.; Jacquemin P., Lannoy V., Rouseau G.G., Lemaigre F.P.; Jacquemin P., Lannoy V., Rouseau G.G., Lemaigre F.P.; Jacquemin P., Lannoy V., Rouseau G. the ONECUT class of homeodomain transcription factors whose function in liver partially overlaps with that of hepatocyte nuclear factor-6."; J. Biol. Chem. 274:2665-2671 (1999).

-: FUNCTION: TRANSCRIPTIONAL ACTIVATOR. ACTIVATES THE TRANSCRIPTION OF A NUMBER OF LIVER GENES SUCH AS HNF3B.

-: SUBCELLULAR LOCATION: Nuclear.
-: SIMILARITY: BELONGS TO THE CUT HOMEOBOX FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16-OCT-2001 (Rel. 40, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
One cut domain family member 2 (ONECUT-2 transcription factor) (OC-2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens (Human).
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                         PERMI, PD0, 2170, CUT; 1.

PERM, PF00046; homeobox; 1.

Proform; PD000010; Homeobox; 1.

Proform; PS00027; HOMEOBOX; 1.

PROSITE; PS00077; HOMEOBOX 1; FALSE_NEG.

PROSITE; PS00071; HOMEOBOX 2; 1.

Transcription regulation; Homeobox; DNA-binding; Nuclear protein; Activator; Alternative splicing.

DNA_BIND 383 444 HOMEOBOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                     POLY-SER.
A -> AESAMGGSVPSLRITSGGPQLSVPPLP (in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 87; DB 1; Length 465
Pred. No. 0.0042;
2; Mismatches 4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /FTId=VSP_002312.
51067 MW; BD651267FD7AC896 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   485 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           isoform Beta)
                                                                                                                                                                                                                                                                                                                        CUT.
HOMEOBOX.
POLY-HIS.
POLY-HIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE=99115605; PubMed=9915796;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16-0CT-2001 (Rel. 40, Created)
16-0CT-2001 (Rel. 40, Last seqn
28-FEB-2003 (Rel. 41, Last anno
InterPro; IPR007108; Cut homeo.
InterPro; IPR003350; Hmoeo CUT.
InterPro; IPR001356; Homeobox.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                || ||| ||| |||::
124 HHHHHHHHHHPHHQR 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; Y18198; CAB38253.1; -.
TRANSFAC; T03259; -.
Genew; HGNC:8139; ONECUT2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 ННКНННККНННКК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 64.7 hes 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              465 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     604894; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ONC2_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ONECUT2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARSPLIC
```

Sanchez A.,

```
Local Similarity 57.9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 303 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               42
68
80
186
79
                                                                                                                                                                                                                                                                                                                                               SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOX2_MOUSE
ID MOX2_MOUSE
AC P32443;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 9
              8
                                                                                                                                                                                                                                                                                                                                                                                                                               ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NOAZ NOBLA.
PS0222; OSPL6;
01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
HOMEODOX protein MOX-2 (Mesenchyme homeobox 2) (Growth arrest-specific homeobox).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=22388257; PubMed=12477932; Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altachul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Maruslana K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C., Broas S.A., McEwan N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan R.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Lepage D.F., Walsh K., "Molecular cloning and localization of the human GAX gene to 7p21."; Genomics 24:535-540(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Grigoriou M., Kastrinaki M.-C., Modi W., Theodorakis K., Mankoo B.,
                                                                                                                                                               Transcription regulation; Homeobox; DNA-binding; Nuclear protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                               ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pachnis V., Karagogeos D.; "Isolation of the human MOX2 homeobox gene and localization to chromosome 7p22.1-21.3"; Genomics 26:550-555(1995).
                                                                                                                                                                                                                                                                                                                                                                                Score 86.5; DB 1; Length 485;
Pred. No. 0.0049;
2; Mismatches 5; Indels 1
                                                                                                                                                                                                                                                                                                                                             AF21E052EFBE5DA1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           303 AA
                                                                                                                        PROSITE; PS00027; HOMEOBOX 1; FALSE NEG.
PROSITE; PS50071; HOMEOBOX 2; 1.
                                                                                                                                                                                                      CUT.
HOMEOBOX.
POLY-GLY.
POLY-PRO.
                                                                                                                                                                                                                                                                                    POLY-ALA.
POLY-HIS.
                                                                                                                                                                                                                                                                                                                         POLY-SER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         147 КЕНИРИРНИНРИЧНИНОВ 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 ККНН-НКНИНККНЕНККК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE=Embryo;
MEDLINE=95331791; PubMed=7607679;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE=95229154; PubMed=7713505;
InterPro; IPR001350; Hmoeo CUT
InterPro; IPR001356; Homeobox.
                                                                                   ProDom; PD000010; Homeobox; 1.
SMART; SM00389; HOX; 1.
                                                                                                                                                                                                                                                                                                                                               52482 MW;
                                                                                                                                                                                                                                                                                                                                                                                    63.6%;
61.9%;
                                               Pfam; PF02376; CUT; 1.
Pfam; PF00046; homeobox; 1.
                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 61.9°
Matches 13; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEOX2 OR MOX2 OR GAX.
Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                             485 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FISSUE-Heart
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HUMAN
                                                                                                                                                                                Activator
                                                                                                                                                                                                        DNA_BIND
DNA_BIND
                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOX2_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 8
          ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HID
DDT TOD
DD
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseaiseb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                             Cordes M., Lacy M.;
Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
-1-FUNCTION: ROLE IN MESODERM INDUCTION AND ITS EARLIEST REGIONAL SPECIFICATION. SOMITOGENESIS, AND MYCOENIC AND SCLEROTOWAL DIFFERENTIATION. MAY HAVE A REGULATORY ROLE WHEN QUIESCENT VASCULAR SMOOTH MUSCLE CELLS REENTER THE CELL CYCLE (BY
Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Gohnerch A., Schein J.E., Jones S.J.M., Marra M.A.; Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMART; SN00389; HOX; 1.
PROSITE; PS00027; HOMEOBOX 1; 1.
PROSITE; PS50071; HOMEOBOX 2; 1.
Homeobox; DNA-binding; Nuclear protein; Developmental protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 86; DB 1; Length 303;
Pred. No. 0.0036;
3; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      79 POLY-HIS.

85 POLY-GLN.

245 HOMEOBOX.

56 G - D (IN REF. 2).

79 MISSING (IN REF. 2).

33457 MW; 809ADEOCD090023D CRC64;
                                                                                                                                                                                                                                                          Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUBCELLULAR LOCATION: Nuclear (Potential).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TISSUE SPECIFICITY: EMBRYO AND PLACENTA. SIMILARITY: Contains 1 homeobox domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          303 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR; A55641; A55641.
PIR; A56837; A56837.
PIRSP; P14653; B192.
TRANSEAC; T04005; -
Genew; HGNC:7014; MEOX2.
MIM; 600535; -
GO; GO:0008015; P:circulation; TAS.
GO; GO:0008015; P:development; TAS.
InterPro; IPR001356; Homeobox.
InterPro; IPR001356; HTH_lambrepressr.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLY-SER.
POLY-HIS.
POLY-GLN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 КНИНИНИКИНИКИ 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL, X82629, CAA57949.1; -.
EMBL, 136328; AAA58497.1; -.
EMBL; BC017021; AAH17021.1; -.
EMBL; AC004482; AAC06184.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF00046; homeobox; 1.
PRINTS; PR00031; HTHREPRESSR.
ProDom; PD000010; Homeobox; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-OCT-1993 (Rel. 27, Created)
                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 230-303 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      63.2%;
```

```
[2]
REVISIONS.
                                                                                                                   63
                                                                                                                                                                                                     MOX2_RAT
ID _MOX2_RAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 cells."
                                                                                                                                                                                                                                       P39020
                                                                                                                                                                               RESULT 10
                                                                                                                                                                                                                        à
                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBJOURNEE OF 1-11 FROM N.A. (3) SEQUENCE OF 1-11 FROM N.A. (3) SEQUENCE OF 1-11 FROM N.A. (4) Malsh K.;

Andres V., Fisher S., Wearsch P., Walsh K.;

In positive factors including myoryte-specific enhancer factor 2.";

Mol. Cell. Biol. 15:4272-4281(1995).

In positive factors including myoryte-specific enhancer factor 2.";

Mol. Cell. Biol. 15:4272-4281(1995).

In positive factors including myoryte-specific enhancer factor 2.";

Mol. Cell. Biol. 15:4272-4281(1995).

In positive factors including myoryte-specific enhancer factor 2.";

Andres V., Fisher S., Wearsch P., Walsh K.;

In positive factors including the scaling including the vertebral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; 
                                                                                                                                                                                                               MEDLINE-92201999; PubMed=1363541; Candia A.F., Hu J., Crosby J., Lalley P.A., Noden D., Nadeau J.H., Wright C.V.E.; Hu J., Crosby J., Lalley P.A., Noden D., Nadeau J.H., Wright C.V.E.; define a novel homeobox gene subfamily and are differentially expressed during early mesodermal patterning in mouse
                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A. Complete 1909944; C.V.E.; Candia A.F., Kovalik J.-P., Wright C.V.E.; Kovalik J.-P., Wright C.W.E.; Maino acid sequence of Mox-2 and comparison to its Xenopus and rat
                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLY-GLN.
GLN/HIS-RICH (OPA-REPEAT)
HOMEOBOX.
41BD05FC39AA4427 CRC64;
01-OCT-1993 (Rel. 27, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
HOMEODOX protein MOX-2 (Mesenchyme homeobox 2)
MEOX2 OR MOX2 OR MOX-2 OR GAX.
Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   and costal precursors.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nucleic Acids Res. 21:4982-4982(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; Z16406; CAA78899.1; -.
EMBL; S79168; -; NOT_ANNOTATED_CDS.
PIR; B49122; B49122.
                                                                                                                                                                                                                                                                                                                                 Development 116:1123-1136(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 245 H
33506 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 186 2
303 AA;
                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA BIND
SEQUENCE
                                                                                                                                                                                                                                                                                                                             embryos.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      homologs.
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Molecular cloning of a diverged homeobox gene that is rapidly down-regulated during the G0/G1 transition in vascular smooth muscle
                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SUBJECT (MAR-1993) to the EMBL/GenBank/DDBJ databases.
-1- FUNCTION: ROLE IN MESODERM INDUCTION AND ITS EARLIEST REGIONAL
SPECIFICATION, SOMITOGENESIS, AND MYGGENIC AND SCIEROTOWAL
DIFFERENTIATION: MAY HAVE A REGULATORY ROLE WHEN QUIESCENT
VASCULAR SMOOTH MUSCLE CELLS REENTER THE CELL CYCLE.
-1- SUBSCELLULAR LOCATION: Nuclear (Potential).
-1- TISSUE SPECIFICITY: AORTA AND HEART. ALSO DETECTED IN LUNG AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TISSUE=Aorta;
MEDLINE=93268321; PubMed=8098844;
Gorski D.H., Lepage D.F., Patel C.V., Copeland N.G., Jenkins N.A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -1- INDUCTION: RAPIDLY AND TRANSIENTLY DOWN-REGULATED DURING THE TRANSITION FROM G0 TO G1 INDUCED BY MITOGEN STIMULATION.
-1- SIMILARITY: Contains 1 homeobox domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Developmental protein.
                                                                   ö
Score 86; DB 1; Length 303;
Pred. No. 0.0036;
3; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                               01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
28-FEB-2003 (Rel. 41, Last amnotation update)
Homeobox protein MOX-2 (Growth arrest-specific homeobox).
MEOXZ OR MOXZ OR MOX-2 OR GAX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLY-HIS.
POLY-GLN.
GLN/HIS-RICH (OPA-REPEAT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterProj irrucce...

Pram. PF00046; homeobox; 1.

PRINTS; PR00021; HTHREPRESSR.

PRINTS; PR00031; HTHREPRESSR.

PRODOM; PD00010; HOmeobox; 1.

SWART; SM00389; HOX; 1.

PROSITE; PS00027; HOMEOBOX_1; 1.

PROSITE; PS00071; HOMEOBOX_2; 1.

PROSITE; PS50071; HOMEOBOX_2; 1.

PROSITE; PS50071; HOMEOBOX_2; 1.

PROMEOBOX; DNA-binding; Nuclear protein; Dev
                                                                                                                                                                                                                                                                                                                                                                303 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001356; Homeobox.
InterPro; IPR000047; HTH lambrepressr.
PF00046; homeobox; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mol. Ceil. Biol. 13:3722-3733(1993).
                                                                                                                                      2 КНИНКНИНККИННКК 20
                                                                                                                                                                                     EMBL, Z17223; CAA78931.1; -.
PIR; A48130; A48130.
HSSP, P14653; 1B72.
TRANSFAC; T04054; -.
             63.2%;
                    Query Match
Best Local Similarity 57.9
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
```

```
m
                                                                                      NR43 HUMAN
                                    셤
               ઠે
                                                             .
.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                   01-NOV-1997 (Rel. 35, Created)
1-NOV-1997 (Rel. 35, Last sequence update)
1-NOV-1997 (Rel. 35, Last annotation update)
Nuclear hormone receptor NOR-2 (Neuron-derived orphan receptor 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF00105; zf-C4; 1.—
PRINTS; PR00047; STROIDFINGER.
PRODOM; PD0000035; Znf_C4steroid; 1.
SMART; SR00399; Znf_C4; 1.
PROSITE; PS00031; NUCLEAR RECEPTOR; 1.
Receptor; Transcription regulation; DNA-binding; Nuclear protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7
                                                              ö
                            63.2%; Score 86; DB 1; Lengun ...,
57.9%; Pred. No. 0.0036;
...marches 5; Indels .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 430;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 85; DB 1; Length 430
Pred. No. 0.0064;
2; Mismatches 2; Indels
             303 AA; 33605 MW; 7776642AEFA3A2E8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A30E42899EF60FFB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLY-HIS.
POLY-GLN.
POLY-ALA.
POLY-ALA.
POLY-ALA.
POLY-ALA.
                                                                                                                                                                             430 AA.
HOMEOBOX
                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
STRAIN-Sprague-Dawley; TISSUE-Brain;
MEDLINE-96000221; PubMed-1556683;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HSSP; P19793; ZNLL.
TRANSFAC; T04752; -.
InterPro; IPR001628; Znf_C4steroid.
                                                                                                                                                                             PRT;
                                                                                      2 КНИНКИНИККИНИКИНКК 20
                                                                                                     :|| ||| ||| ::
63 QHHRGHHHHHHHHQQ 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    46297 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            62.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; X86003; CAA59993.1; -. PIR; S66671; S66671.
                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 66.7
Matches 12; Conservative
                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                  Rattus norvegicus (Rat).
  245
                                                 Local Similarity
hes 11; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    430 AA;
                                                                                                                                                                                                                                                                                                     NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         subfamily
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Zinc-finger.
 DNA BIND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ZN FING
ZN FING
SEQUENCE
                                      Query Match
                                                                                                                                                                             NOR2 RAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                        <u>06351</u>6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                             Matches
                                                                                      ઠ
                                                                                                            셤
```

```
Labelle Y., Zucman J., Stemman G., Kindblom L.-G., Knight J., Turc-Carel C., Dockhorn-Dworniczak B., Mandahl N., Desmaze C., Perer M., Aurias A., Delattre O., Thomas G. "Oncogenic conversion of a novel orphan nuclear receptor by chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 1-69 AND 301-443 FROM N.A. (ISOFORM BETA), AND ALTERNATIVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. (ISOFORM BETA), AND CHROMOSOMAL TRANSLOCATION WITH
                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE=Fetal brain;
MEDLINE=96404972; PubMed=8809112;
Ohkura N., Ito M., Tsukada T., Sasaki K., Yamaguchi K., Miki K.;
Structure, mapping and expression of a human NOR-1 gene, the third member of the Nur77/NGFI-B family.";
Biochim. Biophys. Acta 1308:205-214(1996).
                                                                                                                                                                                                                                                                                                       Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                             orphan receptor 1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE-Skeletal muscle;
MEDLINE=98241504; PubMed=9573341;
Ohkura N., Ito M., Tsukada T., Sasaki K., Yamaguchi K., Miki K.;
"Alternative splicing generates isoforms of human neuron-derived orphan receptor-1 (NOR-1) mRNA.";
Gene 211:79-85(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Clark J., Benjamin H., Gill S., Sidhar S., Goodwin G., Crew J., Gusterson B.A., Shipley J., Cooper C.S.; "Fusion of the EWS gene to CHN, a member of the steroid/thyroid receptor gene superfamily, in a human myxoid chondrosarcoma."; Oncogene 12:229-235(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. (ISOFORM ALPHA).
TISSUE=Peripheral blood;
MEDLINE=9619925; Pubmed=8614405;
Hedvat C.V., Irving S.G.;
"The isolation and characterization of MINOR, a novel mitogen-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IsoId=Q92570-2; Sequence=VSP_003712, VSP_003713;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Event=Alternative splicing; Named isoforms=2;
                                                                                                                             NR43 HUMAN STANDARD; PRT; 626 AA.
092570; 012335; 014979; 016420; 09UEK2; 09UEK3; 01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Nuclear hormone receptor NOR-1 (Neuron-derived c/Mitogen induced nuclear orphan receptor).
NR43 OR NOR1 OR MINOR OR CHN OR CSWF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -1- FUNCTION: BINDS TO THE BIA RESPONSE-ELEMENT.
-1- SUBCELLULAR LOCATION: Nuclear (Potential).
-1- ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Name=Alpha;
IsoId=Q92570-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     inducible nuclear orphan receptor."; Mol. Endocrinol, 9:1692-1700(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hum. Mol. Genet. 4:2219-2226(1995).
                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A. (ISOFORM ALPHA).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A. (ISOFORM ALPHA).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE=Fetal brain;
MEDLINE=96152889; Pubmed=8570200;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE=Fetal heart;
MEDLINE=96177652; PubMed=8634690;
                       99 нинини--ниминиоо 114
нинкнинккинкк 20
                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          translocation.";
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@leb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             YCPTDQAAAGTDAEHVQQFYNLLTA -> VSFMISCFQMND QGLYLWLLYIRVD (in isoform Beta).
/FIId=VSP_003712.
Missing (in isoform Beta).
                                                                                     -!- DISEASE: A FORM OF EXTRASKELETAL MYXOLD CHONDROSARCOMAS (EMC) IS CHARACTERIZED BY A CHROMOSOWAL TRANSLOCATION T(9;17) (Q22;Q11) THAT INVOLVES TAF2N AND NR4A3.
-!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR4
                                                                                                                                                            -i- CAUTION: Ref.2 sequence differs from that shown due to frameshifts in positions 4 and 20.
-!- TISSUE SPECIFICITY: High expression of isoform alpha in skeletal muscle. High expression of isoform beta in skeletal muscle and low expression in fetal brain and placenta.
-!- DISEASE: EWING'S SARCOMA IS CHARACTERIZED BY CHROWOSOMAL TRANSLOCATIONS T(9;22) (Q22-31;Q11-12) WHICH INVOLVES EWS AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00399, ZnF_C4; 1.
PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
Receptor; Transcription regulation; DNA-binding; Nuclear protein; Zinc-finger; Chromosomal translocation; Proto-oncogene;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62.5%; Score 85; DB 1; Length 626; ilarity 66.7%; Pred. No. 0.009; Conservative 2; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MIM; 600542; -.
GO; GO: 0005488; F: binding activity; TAS.
GO; GO: 0003707; F: steroid hormone receptor activity; TAS.
GO; GO: 0004877; F: thyroid hormone receptor activity; TAS.
InterPro; IPR00536; Hormone_rec_lig.
InterPro; IPR001723; Stdhrmn_receptor.
InterPro; IPR001628; Znf_C4steroid.
Pfam; PF00104; hormone_rec_l.
Pfam; PF00105; zf-C4; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LIGAND-BINDING (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40D18DA8FEF991EC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLY-SER.
NUCLEAR RECEPTOR-TYPE.
                                                                                                                                                                                                                                                                                                                                    EMBL; D78579; BAA11419.1; -... EMBL; U12767; AAB02581.1; ALT_FRAME. EMBL; S81243; AAB36006.1; ALT_INIT. EMBL; X88894; CAA61984.1; -... EMBL; D85241; BAA28608.1; -... EMBL; D85242; BAA31221.1; -... HSSP; P19793; ZNLL. TRANSFAC; T04750; -... TRANSFAC; T04750; -... TRANSFAC; T04750; -...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLY-ALA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLY-ALA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLY-HIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRINTS; PRO0398; STRDHORMONER.
PRINTS; PRO0047; STROIDFINGER.
PRODOM; PRO00035; Znf_C48teroid; 1.
SMART; SM00430; HOLI; 1.
SMART; SM00399; ZnP_C4; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        68199 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Genew, HGNC:7982; NR4A3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240
454
579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2220
2231
2287
3357
3490
4490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             splicing.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              454
579
585
626 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alternative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT CONFLICT CONFLICT SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DNA BIND
ZN FING
ZN FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /ARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /ARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
DOMAIN
DOMAIN
DOMAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                           Eukaryota; Mēcazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PEGUIO 1 DE CONTROL DE
                                                                                                                                                                                        01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
Nuclear hormone receptor NOR-1 (Neuron-derived orphan receptor 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62.5%; Score 85; DB 1; Length 628; 66.7%; Pred. No. 0.009; tive 2; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LIGAND-BINDING (POTENTIAL).
9AC285D9A65226D9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLY-HIS.
POLY-GLN.
POLY-ALA.
POLY-ALA.
POLY-ALA.
POLY-SER.
POLY-SER.
NUCLEAR RECEPTOR-TYPE.
                                                                                                                                                     628 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D38530; BAA07535.1; --
PIR; JC2493; JC2493.
HSSP; P19793; ZNLL.
IRANSERAC; T02768; --
InterPro; IPR000536; Hormone rec_lig.
InterPro; IPR001723; Stdhrmn_receptor.
InterPro; IPR001723; Stdhrmn_receptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C4-TYPE
                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           68564 MW;
                                95 ННЯНИН--НИНИНИОО 110
3 ННИКНИНИКНИНКК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 66.7
Matches 12, Conservative
                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             116
1189
222
223
223
223
328
359
359
492
492
                                                                                                                                                                                                                                                                                                                Rattus norvegicus (Rat)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                628 AA;
                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                               NR4A3 OR NOR1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ZN FING
ZN FING
DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
DNA BIND
                                                                                                                                                          NR43 RAT
P51179;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                RESULT 13
NR43_RAT
                                                                                                                                                                                 셤
         ò
```

ï

ä

Gaps

.. 7

Similarity

Query Match Best Local S: Matches 12,

12;

```
RATANTHEMESTREADLY;

RATANTHEMESTREADLY;

RADATHEMESTREADLY;

RADATHEMESTREADLY;

RADATHEMESTREADLY;

RADATHEMESTREADLY;

RADATHEMESTREADLY;

RADATHEMESTREADLY;

RADATHEMESTREADLY;

RADATHEMESTREADLY;

RADATHEMESTREADLY;

RADATHON, CANADATHAMEN, Blazel;

RADATHAMESTREADLY;

RADATHAMEST
                                                                                                                                                                                                                                                                                                                                                                MEDLINE=96038094; PubMed=7555710;
Treisman J.E., Lai Z.-C., Rubin G.M.;
"Shortsighted acts in the decapentaplegic pathway in Drosophila eye
development and has homology to a mouse TGF-beta-responsive gene.";
Development 121:2835-2845(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -!- FUNCTION: Probable transcription factor required for peripheral
                                                                                                                 BUN2_DROME STANDARD; PRT; 1211 AA.
024523; 09VK79; 09VK79;
28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Bunched protein, class 2/class 3 isoforms (Shortsighted protein)
BUN OR SHS OR CG5461.
Drosophila melanogaster (Fruit fly).
                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. (ISOFORM CLASS 2), SUBCELLULAR LOCATION, AND
                                                                                                                                                                                                                                     Eukaryota, Metazoa, Arthropoda, Hexapoda, Insecta, Pterygota, Neoptera, Endopterygota, Diptera, Brachycera, Muscomorpha, Ephydroidea, Drosophilidae, Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A. (ISOFORMS CLASS 2 AND CLASS 3).
                            ||| ||| ||| |||::
           нинкиникиникк 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Science 287:2185-2195(2000).
                                                                                                                                                                                                                                                                                                                                                    TISSUE=Eye-antennal disk;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=Berkeley;
                                       66
            ઠે
                                       셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
be required for the transmission of the dpp signal and for a morphogenetic movement of the medulla in the brain that reorients the second optic lobe relative to the first. Plays a role in determining proper dorsal cell fates leading to the formation of SUBCELDULAR LOCATION: Nuclear and cytoplasmic.

ALTERNATIVE PRODUCTS:
                                                                                                                                                                                                                                                                                                                                       IBOId=024523-2; Sequence=VSP 006670; SIMILARITY: BELONGS TO THE TSC-22/DIP/BUN FAMILY. CAUTION: Ref.2 sequence differs from that shown due to erroneous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HSSP; P80220; IDIP.
FlyBase; FBGN0010466; bun.
InterPro; IRR00580; TSC-22_Dip_Bun.
Pfam; PF01166; TSC-22; 1.
Probom; P001125; TSC-22_Dip_Bun; 1.
PROSITE; PS01289; TSC-22_Tip_Bun; 1.
Transcription regulation; Nuclear protein; Alternative splicing.
                                                                                                                                                          Event=Alternative splicing; Named isoforms=3; Comment=Experimental confirmation may be lacking for some
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 85; DB 1; Length 1211;
Pred. No. 0.016;
2; Mismatches 4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /FTIG-VSP 006670.

K -> E (IN REF. 1).
MISSING (IN REF. 1).
Q -> QQQ (IN REF. 1).
QQVTSAA -> TS (IN REF. 2).

MW; 9925A9159A7051B0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (in isoform Class 3)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NE-ZIPPER.
                                                                                                                                                                                                                                                    IsoId=Q24523-1; Sequence=Displayed;
                                                                                                                                                                                                                                                                      Name=Class 1;
IsoId=Q24522-1; Sequence=External;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLY-GLN.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLY-GIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLY-ALA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLY-ALA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLY-ALA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Missing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AE003636; AAF53200.1; ALT_SEQ.
EMBL; AE003636; AAF53201.1; ALT_SEQ.
PIR; T13804; T13804.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62.5%;
ilarity 64.7%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            125299
                                                                                                                                                                                                                                                                                                                                                                                                              gene model prediction.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; L42512; AAC41608.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 947
1005
1014
1090
1201
109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
Matches 11; Conser
                                                                                                                                                                                                                               Name=Class 2;
                                                                                                                                                                                                                                                                                                                       Name=Class 3
                                                                                                                                                                                                             isoforms;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
```

nervous system morphogenesis, eye development and oogenesis. May

3 ННИКНИККНИКИ 19

ò

```
REC STRANSBERGERS PUBMED-10731112;

RAMAND-BERGERS S. Holt R.A., Evans C.A., Gocayne J.D.,

RAMAND-CCINIKER S.E., In P.W. Hoskins R.A., Galle R.F.,

RAMATAIGES P.G., Scherer S.E., Li P.W. Hoskins R.A., Galle R.F.,

RAMATAIGES R.A., Lewis S.E., Kichards S., Ashburner M., Henderson S.N.,

ROUGGE R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RAMEDIJJE, Wortman J.R., Yandell M.D., Zhang O., Chen L.K.,

RAMEDIJJE, RAMEN B.E., RAMATAIRS R.A., Pfeiffer B.D.,

RAMEN R.M., Bencs P.V., Bernal J., Barker E.G., Nelson C.R., Miklos G.L.G.,

RAMEN R.M., Bencs P.V., Bernal B.P., Bandarl D., Bolshakov S.,

Berkova D., Botcher M.R., Bouck J., Brokstein P., Brottier P.,

ROKOVA D., Botcher M., Deuck J., Brokstein P., Brottier P.,

RAMEN R.M., Bencs P.V., Bernal B.P., Bandarl D., Bolshakov S.,

RAMEN R.M., Bencs P.V., Bernal B.P., Bandarl D., Bolshakov S.,

ROGSON K., Doup L.E., Downes M., Dugas R., Center A., Chandra I.,

RAMEN R.J., Bencs P.V., Bench R., Davies P.,

RAMEN R.J., Bencs P.V., Bench R., Davies P.,

RAMEN R.J., Bencs P.V., Bench R., Davies P.,

RAMEN R.J., Bench R., Garry J.M., Cabriellan A.E., Davies P.,

RAMEN R.J., Bench R., Rouled J., J., Wei M., H., Ibegwan C.,

RAMEN R.J., Rausellan A.E., Gary N.S., Gelbart W.M., Glasser K.,

RAMEN R.J., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,

RAMEN R.J., Lewitsky A.A., Li J., Li J., Liang Y., Lin X.,

Liu X., Mattei B., Wolfred C.D., Kraft C., Morris J., Moshrefi A.,

RAMEN R.M., Mory W., Murphy B., Murphy L., Murny D.M., Nelson D.L.,

RAMEN R.M., Ramington K., Sauders R., Sand R.,

RAMEN R.M., Ramington K., Sauders R., Well B., Sand R.,

RAMEN R., Reinigton K., Sauders R., Well R., Sand H.,

RAMEN R., Reinigton K., Sauders R., Well R., Sand R.,

RAMEN R., Reinigton K., Sauders R., Well R., Sand R.,

RAMEN R., Ramington R., Sauders R., Well R., Sand R.,

RAMEN R., Reinigton R., Sauders R., Well R., Sand R.,

RAMEN R., Ramington R., Sauders R., Well R., Sand R.,

RAMEN R., Reinigton R., Sand R., Shon R.,

RAMEN R., Ramington R., Sauders R., Ramingto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FUNCTION: REQUIRED FOR THE SEGMENT-SPECIFIC REPRESSION OF HOMEOTIC SELECTOR GENES. PC MAY BE INVOLVED IN THE STABLE TRANSMISSION OF A DETERMINED STATE BY ITS EFFECTS ON CHROMATIN STRUCTURE. PROMOTES
                                                                                                                                                                                                                                                                     Drosophila melanogaster (Fruit fly).

Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Paro R., Hogness D.S.;
"The Polycomb protein shares a homologous domain with heterochromatin-associated protein of Drosophila.";
Proc. Natl. Acad. Sci. U.S.A. 88:263-267(1991).
                                                                                                                                                               01-MAY-1992 (Rel. 22, Created)
01-MAY-1992 (Rel. 22, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
                                                                                                                        390 AA
                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN=Oregon-R;
MEDLINE=91095442; PubMed=1898775;
301 ННОРНИННИННИЙ 317
                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                         Polycomb protein.
PC OR CG7618.
                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=7227;
                                                                                                                                                 P26017; Q9VP49;
01-MAY-1992 (Re
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                               DROME
```

LOCUS-SPECIFIC CHROMATIN COMPACTION.

```
..
H
                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the ENBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in o way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
SUBCELLULAR LOCATION: Nuclear (By similarity).

DEVELOPMENTAL STAGE: REQUIRED DURING THE ENTIRE LARVAL PERIOD FOR NORMAL ADULT DEVELOPMENT. IT IS FOUND IN ALMOST ALL CELLS AND ITSJUES TROUGHOUT GASTRULATION AND ORRANGENESIS THOUGH AT A MUCH LOWER LEVEL THAN IN BARLY SYNCYTIAL STAGES.

SIMILARITY: Contains 1 chromo domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 390;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5DB24AE4B326C3B9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KKHHHHHHHIKSERNSGRRSESPLTНННН 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 84; DB 1;
Pred. No. 0.0074;
0; Mismatches 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 ККНИНКИНИККН-------ННКИНН 18
                                                                                                                                                                                                                   or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLY-HIS.
                                                                                                                                                                                                                                                                                                                           Flybase; FBgn0003042; Pc.
GO; GO:0016458; P:gene silencing; IGI.
InterPro; IPR000953; Chromo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLY-HIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHROMO.
                                                                                                                                                                                                                                                                                                                                                                                                                            PROSITE; PS00598; CHROMO 1; 1.
PROSITE; PS50013; CHROMO 2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         43976 MW;
                                                                                                                                                                                                                                                                    EMBL; AE003594; AAF51707.1; -. PIR; A38565; A38565.
                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF00385; chromo; 1.
PRINTS; PR00504; CHROMODOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        41.2%;
                                                                                                                                                                                                                                                   EMBL; X55702; CAA39229.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       14; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                             SMART; SM00298; CHROMO;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            390 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                    HSSP, P23197; 1AP0.
TRANSFAC; T00693;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nuclear protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
             ++
           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

Search completed: January 20, 2004, 18:23:56 Job time : 6.66667 secs

Q94736 stomoxys ca QBmxa9 holopneuste QBt029 drosophila QBt3u6 drosophila Q97261 drosophila Q97261 drosophila QB11V0 plasmodium QB11V0 plasmodium QB11V0 plasmodium QB14X7 gallus gall Q95m23 mus musculu Q95m23 mus musculu Q95m23 mus musculu Q95m23 musculu Q94b2 drosophila Q94b2 drosophila Q94b2 drosophila Q95r44 drosophila Q95r44 drosophila Q95r44 drosophila Q95r47 pseudomonas Q94477 pseudomonas Q94477 pseudomonas Q9443 drosophila Q95444 drosophila Q95444 drosophila Q95444 drosophila Q95444 drosophila Q9447 pseudomonas Q9443 drosophila Q9443 drosophila Q9443 drosophila

OM protein

Run on:

Sequence:

Searched:

Database

Result No.

```
Score 98; DB 5; Length 1318;
Pred. No. 0.00032;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1318 AA; 144944 MW; BIDA5CA0427066F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Dictyostelium discoideum (Slime mold).
Bukaryota, Mycetozoa, Dictyosteliida, Dictyostelium.
NCBI_TaxID=44689;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OS5PH4,
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Histidine kinase DhkM (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 1318 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ALIGNMENTS
Q94736
QBRZA9
QBRT3U6
Q91814
Q91814
Q9965E0
Q910XT
Q990XT
Q990XT
Q990XT
Q990XT
Q91CR
Q98CB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          72.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                         NON TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    095PH4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 1
     A COOS ON THE COOR OF THE COOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q95ph4 dictyosteli
Q9sg87 arabidopsis
Q9zqa3 arabidopsis
Q9zqa3 arabidopsis
Q8uvj4 petromyzon
Q26056 plasmodium
Q8mm30 dictyosteli
Q8sxd4 drosophila
Q9sxd3 arabidopsis
Q8i1j9 plasmodium
Q9bsn0 homo sapien
Q9bsn0 homo sapien
Q9bsn0 homo sapien
Q9d1u9 mus musculu
Q9f1u9 mus musculu
                                                                                                                                                           January 20, 2004, 18:20:39; Search time 24.6377 Seconds (without alignments) 209.478 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Description
                         GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                      830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries
                                                                                                                      - protein search, using sw model
                                                                                                                                                                                                                                                             US-10-018-103A-6
136
1 KKHHHKHHHKKHHKK 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9SG87
Q9ZQA3
Q9ZQA3
Q26056
Q8MP30
Q8SXD4
Q9ZW23
Q9ZW23
Q9ZW23
Q9ZW23
Q9BSN0
Q
Q
Q9BSN0
Q
Q
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     pp plant:*
sp_rodent:*
sp_rodent:*
sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                               BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_archea:*
sp_bacteria:*
sp_fungi:*
sp_human:*
sp_invertebrate:*
sp_mamman:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_rvirus:*
sp_bacteriap:*
sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SPTREMBL 23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score
```

Gaps

ö

ò 셤

```
y Match
Local Similarity 68.4%; Pred. No. 0.00019;
se 13; Conservative 1; Mismatches ج. تماماً
                                                                                       Lin X.;
Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databaseB.
Submit AC006282; AAD20147.1; -.
InterPro; IPR000070; Pectinesterase.
Pfam; PF01095; Pectinesterase: 1.
PROSITE: PS00503; PECTINESTERASE 2; 1.
SEQUENCE 407 AA; 44973 MW; DBAB7B8E166D0B42 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 нинкнинккинкк 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ккнинкиникиник 19
                                                                                                                                                                                                                                                                                                                                                  51 кнінінінінініні 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        69.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAR-2002 (TrEMBLrel. 20, 01-MAR-2002 (TrEMBLrel. 20, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
       Nature 402:761-768(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Plasmodium lophurae
                                                 SEQUENCE FROM N.A.
STRAIN=cv. Columbia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity
tes 12; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hoxlw (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NON TER
NON TER
SEQUENCE
thaliana.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                    Н
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       026056;
                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     026056
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q8UVJ4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 5
Q26056
ID Q260
AC Q260
DT 01-N
DT 01-0
DE H1st
OS Plas
                                                                                                                                                                                                                                                                  Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
ઠે
                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
STRAIN-cv. Columbia,
STRAIN-cv. Columbia,
STRAIN-cv. Columbia;
STRAIN-cv. Columbia;
STRAIN-cv. Columbia;
Subject Columbia;
Strain-columbia;

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AT2G36710.
Arabidopsis thaliana (Mouse-ear cress).
Bukaryota, Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                  Arabidopsis thaliana (Mouse-ear cress).
Bukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NGEL TAXID=3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.

SECOLOMOSIA:

SECOLOMOSIA:

Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,

Lin X., Kaul S., Town C.D., Benito M.-I., Barnstead M.E.,

Romanning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,

Bowman C.L., White O., Nierman W.C., Fraser C.M.;

"Arabidopsis thaliana chromosome III BAC TYM13 genomic sequence.";

"Arabidopsis thaliana chromosome III BAC TYM13 genomic sequence.";

SANILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

EMBL; ACO11708; AAF19568.1;

IncerPro; IPRO01841; Znf_ring.

Pfam; PF00097; Zf-C3HG4; 1.

SMART; SM00184; RING; 1...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
     ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 97; DB 10; Length 684;
Pred. No. 0.00023;
0; Mismatches 4; Indels
       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          946203A42A7399B1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                  01-MAY-2000 (TrEMBLrel. 13, Last sequence update) 01-MAR-2003 (TrEMBLrel. 23, Last annotation update) Putative RING zinc finger protein.
                                                                                                                                                                                                                                update)
         5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          407 AA
                                                                                                                                                                               684 AA
           Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Created)
                                                                                                                                                                                                                    01-MAY-2000 (TrEMBLrel. 13, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PROSITE; PS50089; ZF RING 2; 1.
Metal-binding; Zinc; Zinc-finger.
SEQUENCE 684 AA; 76659 MW; 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             518 КНИННИНИННЯ 534
           5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           книнкиникини 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    71.3%;
                                                                       168 HHRRHHHRHHHHHH 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10, 01-OCT-2002 (TrEMBLrel. 22,
                                                 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity 76.5
nes 13; Conservative
                                                 3 НИНКИННККИННКИНН
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
             Conservative
                                                                                                                                                                                 PRELIMINARY;
               12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q9ZQA3
Q9ZQA3;
                                                                                                                                                                                    Q9SG87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 3
                                                                                                                                               RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 D9ZQA3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
```

```
Gaps
                                                                                                           Petromyzon marinus (Sea lamprey).

Bukaryota, Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;

Petromyzontiformes; Petromyzontidae; Petromyzon.
                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 94; DB 13; Length 404; Pred. No. 0.00033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              404 AA; 41691 MW; A01537E167772552 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01.007-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NOV-1996 (TrEMBLrel. 22, Last annotation update)
Histidine-rich protein (Fragment).
                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         140 AA
404 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2; Mismatches
                                                                                                                                                                                                                                                                                                                                                  PERMINE; PRODO46; homeobox; 1.
PRININE; PRODO424; HOMEOBOX.
PRODOM; PRODO101; HOMEOBOX; 1.
SWART; SM00389; HOX; 1.
PROSITE; PSS50071; HOMEOBOX 1; 1.
DNA-binding; Homeobox; Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                   Created)
PRT;
```

```
Patel S., Ph
Celniker S.;
  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                              Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P., Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K., Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.; "Sequence and Analysis of Chromosome 2 of Dictyostelium."; Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
STRAIN=Berkeley;
Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise B.,
                                                              늉
                            SEQUENCE FROM N.A.
MEDLINE=86174893; PubMed=3007981;
Irving D.O., Cross G.A.M., Feder R., Wallach M.;
"Structure and organization of the histidine-rich protein gene Plasmodium lophurae.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Drosophila melanogaster (Fruit fly).
Bukaryota, Wetazoa, Arthropoda; Haxapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
                                                                                                                                    5; Length 140;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 233;
                                                                                                                                                        4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL, AC123513; AAM4363.1; -.
Hypothetical protein.
SEQUENCE 233 AA; 28937 MW; AD4C7A2F86BB14B9 CRC64;
                                                                                                              SEQUENCE 140 AA; 18216 MW; ED05EED8C805B2FA CRC64;
                                                                                                                                                                                                                                                                      01-0CT-2002 (TrEMBLrel. 22, Created)
01-0CT-2002 (TrEMBLrel. 22, Last sequence update)
01-0MR-2003 (TrEMBLrel. 23, Last annotation update)
Hypothetical protein.
Dictyostelium discoideum (Slime mold).
Eukarrota, Mycetozoa, Dictyostelium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                  67.6%; Score 92; DB 5; I 75.0%; Pred. No. 0.00022;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            67.6%; Score 92; DB 5; I 75.0%; Pred. No. 0.00035;
                                                                                                                                                                                                                                                     233 AA.
                                                                                                                                                         Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches
                                                                               Mol. Biochem. Parasitol. 18:223-234(1986).
EMBL; M15317; AAA29616.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-JUN-2002 (TrEMBLrel. 21, Created)
01-JUN-2002 (TrEMBLrel. 21, Last seq.
01-MAR-2003 (TrEMBLrel. 23, Last ann
GR03394p (CG32676 protein).
CG32676 OR CG9725 OR CG9732.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0
                                                                                                                                                         .
0
                                                                                                                                                                             3 нинхиниххиних 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   71 HHHHHHHHHHHH 86
                                                                                                                                                                                                10 нининимаринини 25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 НИНХНИНКИНИКНИН 18
                                                                                                                               Ouery Match
Best Local Similarity 75.09
Matches 12, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 75.0
Matches 12; Conservative
                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                             FROM N.A.
                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=44689;
          NCBI_TaxID=5853;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q8SXD4; Q9W2R7;
                                                                                                                                                                                                                                                                                                                                                                        STRAIN=AX4;
                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                   TER
                                                                                                                                                                                                                                                                08MP30;
                                                                                                                                                                                                                                                     Q8MP30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08SXD4
                                                                                                                                                                                                                               RESULT 6
QBMP30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 7
SYRERERSO
                                                                                                                                                                              ઠે
                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઠે
```

```
Miranda A. Mungall C.J. Nuncol J. Racleb J., Baragas V., Park S., Callikes S. Dentel S., Phounneanavong S., Wan K., Yu C., Levis S.E., Rubin G.M., Callikes S. Phounneanavong S., Wan K., Yu C., Levis S.E., Rubin G.M., Callikes S. Dentel S., Phounneanavong S., Wan K., Yu C., Levis S.E., Rubin G.M., Callikes S. Dentel S., Phounneanavong S., Wan K., Yu C., Levis S.E., Rubin G.M., Storner S.E., Dentel S., Phounneanavong S., Wan K., Yu C., Levis S.E., Rubin G.M., Storner S.E., Dentel S., Change M., Dentel S., Change M., Dentel S. Dentel
```

```
RESULT 10
                                                                                                          RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0813J9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPARENT SERVICE SERVI
                                                                                                                                                              엄
                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

STRAIM-C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;

KRAMI J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

A Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

A Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

A Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

A Aizawa K., Isawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

Radota K., Matsuda H.A., Ashburnar M., Batalov S., Casavant T.,

Radota K., Matsuda H.A., Ashburnar M., Batalov S., Casavant T.,

Ruchl P., Lewis S., Matsuo Y., Nikaido I., Psoole G., Quackenbush J.,

Ruchl P., Lewis S., Matsuo Y., Nikaido I., Psoole G., Quackenbush J.,

Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

Blake J., Boffelli D., Bojunga N., Carninoi P., de Bonaldo M.F.,

Blake J., Boffelli D., Bojunga N., Carninoi P., de Bonaldo M.F.,

Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

Brownstein M.J., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

Nordone P., Hill B., Righand M., Rodriguez I., Sakamoto N.,

Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

Whymbaw Boris A., Yoshida K., Hasegawa Y., Rawaji H., Kohtsuki S.,

Whymshaw Boris A., Yoshida K., Hasegawa Y., Rawaji H., Kohtsuki S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Functional annotation of a full-length mouse cDNA collection.";
Nature 409:685-690(2001).
EMBL; AK014425; BAB29342.1;
MGD; MGI:1917945; 3830408D24Rik.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5
                                                                                     Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J., submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         66.9%; Score 91; DB 11; Length 109; 68.4%; Pred. No. 0.00023; ive 2; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 1122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 67.6%; Score 92; DB 5; Length 112
75.0%; Pred. No. 0.0014;
ive 0; Mismatches 4; Indels
                                                                                                                                                                                                             Subjected (SEP-2002) to the EMBL/GenBank/DDBJ databases.
EMBL; AY094669; AAM11022.1; --
EMBL; AE003451; AAF46623.2; --
FlyBase; FBGN0052676; CG32676.
InterPro; IPR000626; Ubiquitin.
Pfam; PF00240; ubiquitin; 1.
Pfam; PR00213; UBQ; 1.
PROSITE; PS60633; UBIQUITIN.2; 1.
SEQUENCE 1122 AA; 112183 MM; ABFFBB0F79B09B23 CRC64;
"Annotation of Drosophila melanogaster genome.";
Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01AF85E03F763BB0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     109 AA; 12910 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                336 НИНИНИНИНИНИН 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-07N-2001 (TrEMBLrel. 17, 01-07N-2001 (TrEMBLrel. 17, 101-007-2002 (TrEMBLrel. 22, 13830408D24Eik protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 НЯНКИННККИННКИН 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity 68.4 ies 13; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity 75.0 hes 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                          FlyBase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ОЭревэ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
```

```
[1]

SEQUENCE FROM N.A.

STRAIN=Cv. Columbia;

MEDLINE=20083487; PubMed=10617197;

MEDLINE=20083487; PubMed=10617197;

Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,

Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,

Lin X., Kaul S., Rounsley S.D., Barnstead M.E., Feldblyum T.V.,

Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,

Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,

Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,

Copenhaver G.P., Preuss D., Nierman M.C., White O., Eisen J.A.,

Salzberg S.L., Fraser C.M., Venter J.C.;

Sequence and analysis of chromosome 2 of the plant Arabidopsis
                                                                                                                                                                                                                                                                                                                                                                                                                                  AT2G29410.
Arabidopsis thaliana (Mouse-ear cress).
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota, Viridiplantae, Streptophyta; Embryophyta, Tracheophyta;
Spermatophyta: Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
Spermatophyta: Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
Devlin K., Baker S., Davies P., Mungal K., Berriman M., Pain A.,
Hall N., Bowman S., Churcher C., Quail M., Barrell B.;
Submitted (SEP-2002) to the BMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Plasmodium falciparum (isolate 3D7).
Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=36329;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               66.9%; Score 91; DB 10; Length 385; 66.7%; Pred. No. 0.00072; ive 3; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AC004561; AAC95197.1; -: InterPro; IPR003524; Cation efflux.
Pfam; PF01545; Cation efflux; 1.
TIGRFAMS; TIGR01297; CDF; 1.
SEQUENCE 385 AA; 43534 MW; 49A7B564F57FDC2B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q813J9;
01-MAR-2003 (TrEMBLrel. 23, Created)
01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Hypothetical protein.
                                                                                                                                                                                                                                                                                            Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                           385 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              [2]
SEQUENCE FROM N.A.
MEDLINE=22255708; PubMed=12368867;
                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ||| ||: || ||: ||||: 201 HHHHHHNHKHQHQHHHKE 218
                                     ||| || || |: ||||||||
82 HHHHHHSPHRLHHHKHHHR 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 ННИКНИНККИННКК 20
з нинкин--нккиникинк 19
                                                                                                                                                                                                                                                                                                01-MAY-1999 (TrEMBLrel. 10, 01-MAY-1999 (TrEMBLrel. 10, 01-OCT-2002 (TrEMBLrel. 22,
                                                                                                                                                                                                                                                                                                                                                                                                   Putative zinc transporter.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Guery Match
Best Local Similarity 66.7
Matches 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nature 402:761-768(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
STRAIN=cv. Columbia;
                                                                                                                                                                                                                                     Q9ZW23
Q9ZW23;
```

```
451 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-DEC-2001 (TrEMBLrel. 19, Last sequence up
01-OCT-2002 (TrEMBLrel. 22, Last annotation
Dvl-binding protein NKD2 (Naked cuticle-2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          461 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=21490203; PubMed=11604995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Int. J. Oncol. 19:963-969(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ::| | |||: ||| |||
429 ОКНЕННИНЖЕНИНИНИ 446
                                                                                                                                                                                                                                                                                                                                                             1:| | |||: ||| ||| 444
                                                                                                                                                                                                                                                                                                                                    1 ККНИКИНИКИНИИНИ 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 ККИННИЧЕНКИНИНИ 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-JUN-2001 (TrEMBLrel. 17, C)
01-JUN-2001 (TrEMBLrel. 17, Li
01-MAR-2003 (TrEMBLrel. 23, Li
2210403110Rik protein.
NKD2 OR 2210403110Rik.
Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (TrEMBLrel. 19,
                                                                                                                                                                                                                                                                                          11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Genew; HGNC:17046; NKD2
SEQUENCE 451 AA; 500
                                                                                                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE=Fetal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Katoh M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            and NKD2.
                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q9D7U9
Q9D7U9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q969F2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q969F2
                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q969F2
  ò
                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COSCILLATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K., Buckee C.O., Burrows C., Cherevach I., Chillingworth C., Christodoulou Z., Clark R., Corton C., Chrillingworth T., Christodoulou Z., Clark R., Corton C., Cronin A., Davies R., Davies P., Dearden F., Doggett J., Reltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z., Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P., Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P., Mights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N., Line A., Maddison M., Mclean J., Money P., Moule S., Murphy L., Oliver K., Ormond D., Price C., Quail M.A., Rabbinowitsch E., Alayland D., Price C., Quail M.A., Rabbinowitsch E., Araylor K., Sharp S., Smith R., Squares R., Squares S., Stevens K., Araylor K., Tivey A., Unwin L., Whitehead S., Woodward J., Sulston J.E., Craig A., Newbold C., Barrell B.G;
Mature 419:227-531 (2002).

RR BEBL, ALSSESS, CADSIGES, 1, -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Butheria, Primates, Catarrhini, Hominidae, Homo.
  Churcher C., Harris B., Harris D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   65.8%; Score 89.5; DB 5; Length 4524; 63.6%; Pred. No. 0.0098; ive. 4; Mismatches 1; Indels 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                65.4%; Score 89; DB 4; Length 325; 61.1%; Pred. No. 0.0011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                         4524 AA; 525287 MW; 542206173C29567A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Strausberg R.;
Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      325 AA; 36283 MW; 61A24B05707F82FD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-CCT-2002 (TrEMBLrel. 22, Last annotation update)
Similar to RIKEN CDNA 2210403L10 gene (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-WAR-2003 (TrEMBLrel. 23, Created)
01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
1-WAR-2003 (TrEMBLrel. 23, Last annotation update)
Naked cuticle 1 homolog.
Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    325 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 449 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4413 |||:|| || |||:||:|
4413 KHHNKHSKHHSKHHNKHHNKQ 4434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 КИНИКН----НИККИНИКНИКК 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             303 окненининенинин 320
Hall N., Pain A., Berriman M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 ККИННКИННКИННИН 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; BC004940; AAH04940.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity 63.6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                    Hypothetical protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
tes 11; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ISSUE=Uterus;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NON TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q8C4J8
Q8C4J8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9BSN0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9BSN0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ξ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9BSN0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STAR REPRESENTATION OF THE PROPERTY OF THE PRO
$$$$$$$$$$$$$$$$$$$$$$$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ઠે
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
```

```
STRAIN-C57BL/61; TISSUE-Head;
MEDINE-22354681; PubMed=12466851;
The FANTOM Consortium.
The FANTOM Consortium.
The RIKEN Genome Exploration Research Group Phase I & II Team;
The Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length CONAs.";
Nature 420:563-573 (2002).
EMBL; AK081904; BAC38367.1; -.
SEQUENCE 449 AA; 50230 MW; FC6ADFBE8836D671 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                   Gарв
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gape
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            "Molecular cloning, gene structure, and expression analyses of NKD1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDLINE-21255530; PubMed=11356022;
MEDLINE-21555530; PubMed=11356022;
Wharton K.A. Jr., Zimmermann G., Fousset R., Scott M.P.;
"Vertebrate proteins related to Drosophila naked cuticle bind dishevelled and antagonize wnt signaling.";
Dev. Biol. 234:93-106(2001).
EMBL; AB062887; BAN5001.1;
EMBL; AF359137; AAK57486.1;
                                                                                                                                                                                                                                                                                                                                          65.4%; Score 89; DB 11; Length 449; 61.1%; Pred. No. 0.0014; 1. Indels ive 3; Mismatches 4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   65.4%; Score 89; DB 4; Length 451; 61.1%; Pred. No. 0.0014;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    451 AA; 50055 MW; F27E708F9FAD1F2A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
```

```
ô
                                                                     PERGURNEE FKOW N.A.

STRAIN-C57BL/60; TISSUE=Stomach;

X MEDLINE=1086660; PubMed=11217851;

X MEDLINE=21086660; PubMed=11217851;

X Rawai J., Shinaqawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Kawai J., Shinaqawa A., Shibata K., Konno H., Adachi J., Fukuda S., Arakawa T., Eawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Asito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Kuchl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Kuchl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L.M., Stambli F., Suzuki R., Tominci M., Magner L., Washio T., Sakai K., Okido T., Puruno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Barcher C., Fuljita M., Gariboldi M., Loons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakameto N., Choenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Sato K., Wang K.H., Weitz C., Whittaker C., Wilming L., Mynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Functional annotation of a full-length mouse cDNA collection.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
STRAIN=CS7BL/6xCBA; TISSUE=Lung;
MEDLINE=2125530; PubMed=11356022;
Mharton K.A. Jr., Zimmermann G., Rousset R., Scott M.P.;
"Vertebrate proteins related to Drosybhila naked cuticle bind
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65.4%; Score 89; DB 11; Length 461; ilarity 61.1%; Pred. No. 0.0015; Conservative 3; Mismatches 4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 89; DB 11; Length 461;
Pred. No. 0.0015;
3; Mismatches 4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; Indianal InterPro; InterPro; InterPro; IPRO02049; EF-hand.
Prom; PF000036; efhand; 1.
PROSITE; PS00018; EF HAND; 1.
SEQUENCE 461 AA; 51501 MW; 7C3B550024E87D27 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF00036; efhand; 1.
PROSITE; PS00018; EF HAND; 1.
SEOUENCE 461 AA; 51474 MW; 5830BDDA14E868FD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  461 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 65.4%; Score 89; DB Best Local Similarity 61.1%; Pred. No. 0.00 Matches 11; Conservative 3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  dishevelled and antagonize wnt signaling."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ::| | |||: ||| |||
439 ОRНЕННИНЕННИННИ 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dev. Biol. 234:93-106(2001).
EMBL; AF358136; AAK57485.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 ККНИНКИННККИНН 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Nature 409:685-690(2001).
EMBL, AKO08844; BAB25908.1; -.
MGD, MGI:1919543; Nkd2.
InterPro; IPR002048; EF-hand.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-DEC-2001 (TrEMBLrel. 19,
01-DEC-2001 (TrEMBLrel. 19,
01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MGD; MGI:1919543; Nkd2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match
Best Local Similarity
Matches 11; Conserva
                        NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Naked cuticle-2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hayashizaki
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q91Y45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    년
''' '''
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ....
δ·
```

Gaps

ö

```
1 КИННИКНИККИНКИНИ 18
 ò
         g
```

completed: January 20, 2004, 18:26:25 ne : 25.6377 secs Search com